2-1-1976

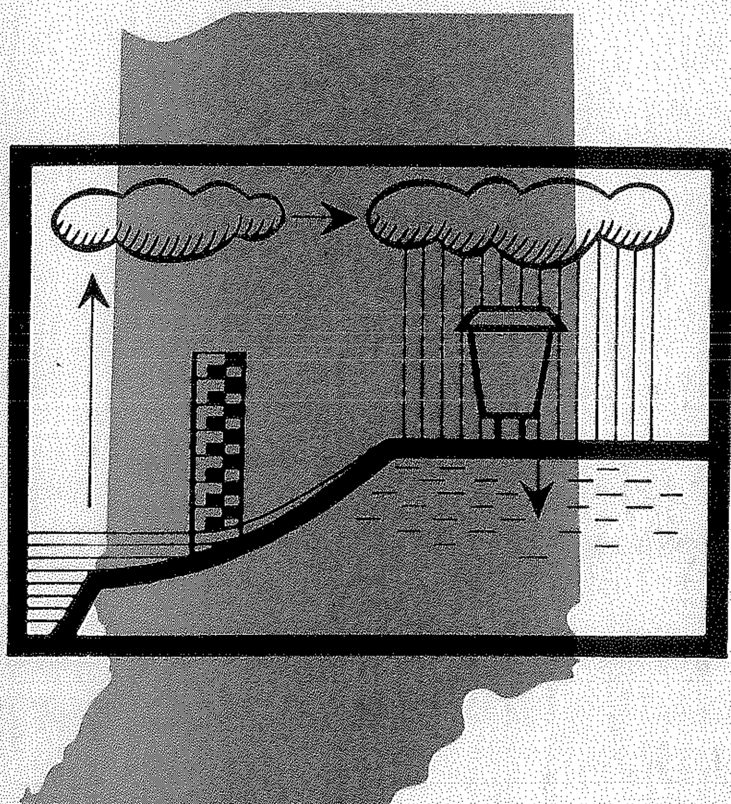
Analysis Of The Effects Of Urbanization On Runoff Characteristics By Nonlinear Rainfall-Runoff Models

R. G.S. Rao

A. R. Rao

TECHNICAL REPORT NO. 58

ANALYSIS OF THE EFFECTS OF URBANIZATION ON RUNOFF CHARACTERISTICS BY NONLINEAR RAINFALL-RUNOFF MODELS



by

R. G. Srinivasa Rao

A. Ramachandra Rao

PURDUE UNIVERSITY
WATER RESOURCES RESEARCH CENTER
WEST LAFAYETTE, INDIANA

Water Resources Research Center
Purdue University
West Lafayette, Indiana

Analysis of the Effects of Urbanization on Runoff Characteristics by Nonlinear Rainfall-Runoff Models

by

R. G. Srinivasa Rao
A. Ramachandra Rao

The work upon which this report is based was supported in part by funds provided by the United States Department of the Interior, Office of Water Research and Technology, as authorized by the Water Resources Research Act of 1964 (P188-379) as amended.

Period of Investigation: September 1971 - August 1975. Partial completion report for OWRT-C-3277.

Grant No. ~~14-31-0001-3713~~

ACKNOWLEDGEMENTS

The work reported herein is based on a part of the doctoral thesis entitled, "Prediction of Daily Runoff and Evaluation of the Effects of Urbanization on Runoff by Using Nonlinear Rainfall-Runoff Models," submitted by R. G. Srinivasa Rao to the Graduate Faculty at Purdue University in December 1975. The work was performed under the supervision of Professor A. Ramachandra Rao. This report covers part of the work performed under the title II project C-3277 entitled, "Systematic Development of Methodologies in Planning Urban Water Resources for Medium Size Communities, Phase I".

The authors wish to express their appreciation to Dr. Dan Wiersma, Director of Water Resources Research Center, and to Dr. J. F. McLaughlin, Head of the School of Civil Engineering at Purdue University for their assistance in the administration of the project.

The authors are grateful to the Office of Water Research and Technology and Purdue University for their support.

We would like to thank Mrs. Ramona Hill for typing this report.

ABSTRACT

The basic objective of the research reported herein is to analyze the effects of urbanization on daily runoff characteristics. The daily rainfall and runoff from Salt Creek watershed near Chicago, Illinois, have been used in the analysis. Two different types of mathematical models, the functional series model and a nonlinear stochastic model have been used to analyze the effects of urbanization on runoff. A new method is proposed for estimation of kernel functions of the functional series models. The application of nonlinear stochastic model to characterize the rainfall-runoff process in which the time derivatives of rainfall sequence are used as additional inputs to the model introduces a new type of model. The role of validation tests in rainfall-runoff models is also emphasized. The models are validated both in prediction and simulation modes and the performance of the models is found satisfactory both in prediction and simulation of runoff.

The effects of urbanization on runoff characteristics such as the histograms, correlograms, mass and flow-duration curves, and extreme values have been analyzed by using both the observed and simulated runoff. All these characteristics have been shown to be significantly affected by urbanization, and the changes brought about by urbanization on these characteristics are discussed quantitatively. Thus the use of the nonlinear models developed in the present study for investigating the effects of urbanization on runoff characteristics has been demonstrated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

A_{nm}	Orthogonal coefficients
CP	Calibration period (water years)
$C_n^{(i)}$	Parameters of nonlinear stochastic model
F_0	Constant response of a model
$F_n(\cdot)$	n^{th} order functional
H_0	Constant response of a model
$H_1(\tau)$	First order kernels
$H_2(\tau_1,\tau_2)$	Second order kernels
$H_n(\tau_1,\ldots,\tau_n)$	n^{th} order kernels
$I(t)$	Observed rainfall value at instant t
$J(\cdot)$	Quadratic error criterion
$K[\cdot]$	Characteristic function of a filter
L	Linear model
L1	First order (linear) functional series model for the data during 1952-53
L2	First order (linear) functional series model for the data during 1963-64
M	Length of memory
MS1	Overall nonlinear stochastic model for the data during 1952-53
MS2	Overall nonlinear stochastic model for the data during 1963-64
ML1	Overall first order functional series model for the data during 1952-53
ML2	Overall first order functional series model for the data during 1963-64
MN1	Overall second order functional series model for the data during 1952-53
MN2	Overall second order functional series model for the data during 1963-64
N	Number of observations
N1	Second order (nonlinear) functional series model for the data during 1952-53
N2	Second order functional series model for the data during 1963-64
NL	Second order nonlinear model
PP	Prediction period (water years)
$P_I(i)$	Probability density function of rainfall
$P_{IZ}(i,z)$	Joint probability density function of rainfall and runoff
$Q(t)$	Model runoff at instant t
$Q_j(t)$	j^{th} component of the model runoff (also output of j^{th} subsystem)

$Q_r(t)$	Regenerated runoff
$Q_p(t)$	Predicted runoff
$\overline{Q}_r$	Mean value of regenerated runoff
$\overline{Q}_p$	Mean value of predicted runoff
R	Number of terms of series expansion
RP	Regeneration period (water years)
R_1	Ratio of error mean to runoff mean
R_2	Ratio of error mean square to runoff mean square
R_3	Correlation coefficient between the observed and model output
$\underset{\sim}{S}$	A positive definite weighting function matrix
S1	Nonlinear stochastic model for the data during 1952-53
S2	Nonlinear stochastic model for the data during 1963-64
$Z(t)$	Observed runoff at instant t
$\overline{Z}$	Mean value of $Z(t)$
$a(t)$, $w(t)$	Residuals or regeneration error from overall models
$\underset{\sim}{a}$	Estimate of parameter vector $\underset{\sim}{\alpha}$
$e_r(t)$	Regeneration error or residuals
$e_p(t)$	Prediction error
$e_i(t)$	Error from the i^{th} subsystem
$F(\cdot)$	Characteristic function of a zero memory system (or the subsystems)
m	Highest order of derivative of rainfall
n	Highest order of nonlinear terms in the model
n_f	Number of first order kernels
n_s	Number of second order kernels
n_t	Total number of parameters
$r(k)$	Autocorrelation coefficient at lag k
$r_{12}(k)$	Cross correlation coefficient at lag k
$\underset{\sim}{\alpha}$	Parameter vector ($1xn_t$) in a functional series model
$\underset{\sim}{\beta}$	Parameter vector (1xp) in an error model
$\gamma(k)$	Autocovariance function at lag k
$\gamma_{12}(k)$	Cross covariance function at lag k
$\sigma_{\underset{\sim}{a}}^2$	Variance of the parameter estimates
σ_Z^2	Variance of $Z(t)$

ρ	Variance of residuals (or regeneration errors)
ξ_m	Orthogonal functions
θ_n	Orthogonal functions
$\eta(t)$	Noise sequence
η_j	Class η_j nonlinear system
$\eta_j(t)$	j^{th} component of noise vector (px1)
$\underset{\sim}{\eta}(t)$	Noise vector (px1)
$\overline{}$	Expectation operator
T	Transpose of a matrix or vector
$\lVert\ \rVert^2$	Norm of vector
$\sim$	Matrix or a vector

CHAPTER I

INTRODUCTION AND STATEMENT OF OBJECTIVES

1.1 Introduction

It is well known that urbanization of watersheds changes the characteristics of runoff from these watersheds. Engineers and planners are frequently faced with the problem of estimating the watershed response after the watershed is urbanized. This problem appears in the design of drainage structures in a watershed while it is being planned to be urbanized. Several methods have been proposed to estimate the effect of urbanization of watersheds on runoff.

Savini and Kammerer (1961) were among the first to discuss the effects of urbanization on watershed response, although the emphasis of their work was on the effect of urbanization on regional water resources. *Carter (1961)*, *Anderson (1968)*, *Wiitala (1961)*, *Van Sickle (1962)* and others have investigated the variability in time lag brought about by urbanization. The variability in the unit hydrograph characteristics due to urbanization has been investigated by *Espey et al. (1965, 1966)* and by *Riley and Dhruvanarayana (1968)*. *Rao et al. (1971)* have developed quasi linear models based on the instantaneous unit hydrographs to estimate the effect of urbanization on runoff hydrographs.

The effect of urbanization on peak discharge has been investigated by *Carter (1961)*, *Anderson (1968)*, and by *Stall and Smith (1961)*. All of these investigators found that the urbanization increased the peak discharge and decreased the time to peak. All the analyses were based on storm hydrographs.

Recently, *Speiker (1971)* has investigated the effects of urbanization on stream flow characteristics by using flow - duration curves. *Bras and Perkins (1975)* have used a conceptual model to analyze the effects of urbanization on catchment response. The Stanford watershed model and its variants have also been used to analyze the effects of urbanization on runoff.

Perhaps the most significant conclusion which can be arrived at after a review of the literature on the studies of effects of urbanization on watershed response is that most of the studies are based on analysis of storms. Although the hydrograph methods are very useful in drainage design and hence have received considerable attention, models are needed to estimate the continuous, and not only the storm, hydrographs. These models are of use, for example, in the design of retention basins. Also, if the estimation procedures are recursive, they can be used for optimal control of drainage systems. Finally, the effect of urbanization on the characteristics of stream flows, such as the daily flows, is also of considerable interest but has not been investigated in depth.

Two major types of rainfall-runoff models appear in the literature in urban hydrology, and these are the conceptual and black box models. The conceptual models are based on the physical aspects of the process such as storage and conveyance elements. Black box models, on the other hand, are formulated by considering only the characteristics of observed rainfall-runoff process. Nonlinear functional series models have been used for modeling rainfall-runoff process. The processes in these models are usually assumed to be time invariant and the rainfall to be a lumped variable. Several methods of estimation of parameters in a nonlinear functional model are available in literature and these are briefly summarized in Sec. 1.2. These methods vary in their complexity of estimation of parameters and prediction performances. A comparative study of some of these methods has been reported by *Rao and Rao (1974)*. One of the conclusions of this study was that simpler methods of estimation of parameters of nonlinear functional series model are necessary and that the models obtained from such estimation methods must be tested for their prediction performance also.

Most of the models proposed in literature use only the rainfall values as input to the model. The runoff values, depend not only upon the magnitude of the rainfall, but also on the time derivatives of rainfall. Therefore the rainfall derivative are also important in the models of the rainfall-runoff process. Consideration of the rainfall derivatives is particularly important in urban areas because of rapid basin responses. Differential equation models of rainfall-runoff process in which up to two or three derivatives of rainfall are also used as input have been proposed in literature (*Kulandaiswamy (1964)*). However, these models are linear and have been used only for storm analyses. The drawbacks of these models are the same as of other linear models of the rainfall-runoff process. Therefore, new models of modeling the rainfall-runoff process in which the rainfall derivatives are also used as inputs are necessary.

In view of these considerations the objectives of the present study are as follows.

1.1.1 Statement of Objectives

(1) Develop a simpler estimation procedure to be used with nonlinear functional models of the rainfall-runoff process and test the efficiency of the procedure in regeneration, prediction and simulation modes.

(2) Develop a nonlinear model of the rainfall-runoff process which involves rainfall and its time derivatives and test the model in regeneration, prediction and simulation modes.

(3) Investigate the capability of the models developed in objectives (1) and (2) above to estimate the runoff from watersheds which are urbanized to different degrees, and hence to study the effects of urbanization on runoff.

1.1.2 Organization of the Report

The report is organized as follows. In the remaining portions of this chapter some aspects of lumped linear and nonlinear models and their application to the rainfall-runoff process are discussed. The data used in the study and their characteristics are discussed in Chapter II. In Chapter III a recursive algorithm is proposed for estimation of Kernels in the nonlinear functional models. A nonlinear stochastic model is proposed for the rainfall-runoff process in Chapter IV and the estimation procedure is detailed. These models are applied to the daily rainfall-runoff sequences from the Salt Creek watershed and the capability of the models to characterize the process are investigated in Chapter V. Quantitative results about the effects of urbanization on runoff are also presented in Chapter V. A set of conclusions are given in Chapter VI.

1.2 The Nonlinear Functional Model of the Rainfall-Runoff Process:

1.2.1 General aspects:

The general functional representation of rainfall-runoff process can be written as in eq. 1.1 by assuming lumped rainfall and invariance of watershed response. If $F_i(I(t))$ represents integral functionals of order i as given in eq. 1.2, then the model given in eq. 1.1 will be the same as the voltera functional series.

$$Q(t)=F_0+F_1[I(t)] + F_2[I(t)]+\ldots+F_n[I(t)] +\ldots \qquad (1.1)$$

In eq. 1.1, $Q(t)$ is total response of the system, $I(t)$ is the total rainfall (input), F_0 is the constant response, $F_1[I(t)]$ is the response of a linear system, $F_2[I(t)]$ is the response of a second order system and so on. Thus the total response $Q(t)$ is a sum of a constant response, linear, second and higher order responses. For a physically realizable system the various functionals in eq. 1.1 can be approximated as in eq. 1.2.

$$F_1[I(t)] = \int_0^\infty H_1(\tau)\ I(t-\tau)\ d\tau;\ H_1(\tau) = 0\ \forall\ \tau<0$$

$$F_2[I(t)] = \int_0^\infty \int_0^\infty H_2(\tau_1,\tau_2)I(t-\tau_1)I(t-\tau_2)d\tau_1 d\tau_2;$$

$$H_2(\tau_1,\tau_2) = 0\ \forall\tau_1,\tau_2<0 \qquad (1.2)$$

$$\vdots$$

$$F_n[I(t)] = \int_0^\infty \int_0^\infty \ldots \int_0^\infty H_n(\tau_1,\tau_2,\ldots,\tau_n)I(t-\tau_1)\ \ldots\ I(t-\tau_n)d\tau_1 d\tau_2\ \ldots$$

$$d\tau_n;\ H_2(\tau_1,\tau_2,\ldots\tau_n)=0,\ \forall\ \tau_1,\tau_2,\ldots,\tau_n<0$$

In eq. 1.2, $H_1(\tau)$, $H_2(\tau_1,\tau_2)$, etc. are the kernels associated with respective linear, second order, and higher order responses of the model. Substituting eq. 1.2 in eq. 1.1, we have

$$Q(t)=F_0+\int_0^\infty H_1(\tau)I(t-\tau)d\tau+\int_0^\infty\int_0^\infty H_2(\tau_1,\tau_2)I(t-\tau_1)I(t-\tau_2)d\tau_1 d\tau_2$$

$$+\int_0^\infty\int_0^\infty\cdots\int_0^\infty H_n(\tau_1,\tau_2,\ldots,\tau_n)I(t-\tau_1)I(t-\tau_2)\ldots I(t-\tau_n)d\tau_1 d\tau_2\ldots d\tau_n \quad (1.3a)$$

$$\text{or } Q(t)=\sum_{r=1}^{n}\int_0^\infty\ldots_{(n)}\int_0^\infty H_r(\tau_1,\ldots,\tau_r)I(t-\tau_1)\ldots I(t-\tau_r)d\tau_1\ldots d\tau_r \quad (1.3b)$$

Equation 1.3 is the n^{th} order version of convolution integral for linear system. The discrete version of eq. 1.3 can be written as in eq. 1.4, where M is the memory of the system.

$$Q(t)=F_0+\sum_{\tau=0}^{M}H_1(\tau)I(t-\tau)+\sum_{\tau_1=0}^{M}\sum_{\tau_2=0}^{M}H_2(\tau_1,\tau_2)I(t-\tau_1)I(t-\tau_2)$$

$$+\sum_{\tau_1=0}^{M}\sum_{\tau_2=0}^{M}\cdots\sum_{\tau_n=0}^{M}H_n(\tau_1,\tau_2,\ldots\tau_n)I(t-\tau_1)I(t-\tau_2)\ldots I(t-\tau_n) \quad (1.4)$$

1.2.2 <u>Classes of Nonlinear Systems</u>:

Nonlinear single input - single output systems can be classified into several classes. These classes are designated class η_1, class η_2, etc. by *Zadeh (1953)*, such that each class in the sequence includes all the preceeding classes. The output of a class η_1 nonlinear system can be written as

$$Q(t) = \int_0^\infty K[I(t-\tau),\tau]\, d\tau, \quad (1.5)$$

where $K[I(t-\tau),\tau]$ is called the characteristic function of the class η_1 system. $K[I(t-\tau),\tau]$ may be a nonlinear functional operating on inputs $I(t-\tau)$. Similarly the output of a class η_2 nonlinear system can be written as

$$Q(t) = \int_0^\infty\int_0^\infty K[I(t-\tau_1),I(t-\tau_2),\ \tau_1,\tau_2]\ d\tau_1 d\tau_2, \quad (1.6)$$

where $K[I(t-\tau_1),I(t-\tau_2),\tau_1,\tau_2]$ is a nonlinear function operating on inputs $I(t-\tau_1)$ and $I(t-\tau_2)$ and is called the characteristic function of class η_2 nonlinear system. Similarly we can write the output equations for a class η_n system.

Class η_1 system includes linear systems if the characteristic function is the impulse response of a linear system. It also includes zero memory systems in which the output is the instantaneous function of input. The output in a zero memory system may be written as

$$Q(t) = f[I(t)], \quad (1.7)$$

where $f(\cdot)$ is the characteristic function of zero memory system. The characteristic function $f(\cdot)$ in a class η_1 system may be a linear or a nonlinear functional. If the functional $f(\cdot)$ is a polynomial functional of order n, then such a nonlinear system belongs to class n_n. The second order functional series models of the rainfall-runoff process are discussed in Chapter III. The zero memory nonlinear systems which belongs to class n_1 have been proposed to model the rainfall-runoff process and are discussed in Chapter IV.

1.2.3 Methods of Modelling Rainfall-Runoff Process:

The original mathematical treatment of the theory of nonlinear models was introduced by *Wiener* (1942, 1958) and elaborated on by *Bose* (1956), *Lee and Schetzen* (1956), *Brilliant* (1958), and *George* (1959). A number of investigators have recognized the importance of nonlinear system models in modelling the rainfall-runoff process and have developed methods to estimate the kernels. In several of these methods, the starting point is the Volterra functional series. In some of these methods the unknown kernels in the functional series model are estimated whereas in others the input or the kernels are approximated by polynomials or transformed into new set of variables. In those methods in which the inputs and kernel functions are transformed, predictor equations with a set of unknown coefficients are derived and methods are developed to estimate these coefficients by using the observed rainfall and runoff sequences.

Brendstetter and Amorocho (1970) used meixner polynomials to approximate the kernels while *Jacoby* (1966) used Laguerre polynomial expansion of the inputs. The coefficients of the predictor equation are estimated by using multiple regression procedure in these methods. *Hino et al.* (1971) used two known delay systems whose kernels are approximated by dirac delta functions. The first and second order kernels were then computed by using the correlation technique of *Lee and Schetzen* (1956). *Kuchment and Borschevsky* (1971) worked with continuous signals and used an analog network to compute the coefficients in the predictor equation. Normalized Laguerre networks were used to approximate the kernels and the predictor coefficients were estimated by using an optimization procedure based upon the gradient method. *Bidwell* (1971) used grouped variables of input and estimated the predictor coefficients by using stepwise regression procedure. All those methods are designed for estimation of unknown coefficients of a single input - single output series.

Nonlinear models of rainfall-runoff process with multiple inputs and multiple outputs have also been proposed in the literature. *Harder and Zand* (1969) formulated a two-input model and used stepwise regression procedure to estimate the predictor coefficients. *Boneh and Diskin* (1971) considered a second order Volterra series model and derived a set of equations which may be used for analysis of multiple storms. The mass conservation and nonnegativity output requirements of the rainfall-runoff process were included in this method as constraints on the kernels. The resulting set of equations with constraints were solved by using the *Fletcher-Powell* optimization scheme to estimate the kernels.

CHAPTER II

DATA USED IN THE STUDY AND THEIR CHARACTERISTICS

.1 Data Used in the Study

Daily rainfall-runoff sequences in Salt Creek basin near Arlington Heights, Illinois were used to nvestigate the effects of urbanization on runoff characteristics. Salt Creek basin, located in Cook nd Dupage Counties of the Chicago metropolitan area has a drainage area of 150 sq. mi. and is approxi- ately 35 miles in length. In 1964, the population in the Salt Creek basin was about 400,000 and about 0 percent of the land was urbanized while the population and urbanized land prior to 1956 were about 50,000 and 22 percent respectively. The population in Salt Creek basin is expected to increase to about 50,000 to 650,000 by 1990 with about 70 percent of the land urbanized. The water and land use problems f Salt Creek basin are typical of those of the areas which are located near large urban centers. hese problems are discussed for the Salt Creek basin by *Speiker* (1970).

The mean annual precipitation over the Salt Creek basin is about 30 inches and varied from 22-46 nches during 1871-1966. Most of the rainfall results from thunderstorms of relatively short duration nd high intensity. The groundwater pumping which is the main source of public water supply affects both ow and high flows in the Salt Creek. The low flows are also influenced by effluent discharges into the alt Creek. The details of the characteristics of the Salt Creek basin can be found in *Speiker* (1970).

Speiker has also shown that urbanization in the Salt Creek basin has affected the low and flood flow haracteristics. Intense urbanization of the basin started in 1956. Consequently two sets of two years f daily rainfall-runoff data, corresponding to water years 1952-53 and 1963-64 were used for the analysis f the effect of urbanization on runoff characteristics. The data from the 1952-53 period represents he watershed response in the period when the urbanization was small and those during 1963-64 represents he response in a period when urbanization had increased. Daily rainfall values at Elgin, Illinois and aily runoff values from Salt Creek basin, near Arlington Heights, Illinois were used in the present study. he watershed area of the Salt Creek basin at Arlginton Heights is about 33.7 mi^2. The details of loca- ion of these stations are given in Table 2.1. The mean annual rainfall in these two periods was pproximately equal. The daily rainfall and runoff sequences in these two periods are shown in Fig. 2.1. he rainfall sequence appears to be random in both the periods with no apparent seasonalities. The aximum daily rainfall was 4.47 in. in the first period whereas it was 2.78 in. in the second period. he runoff sequences in the two periods seem to have some seasonality in them. A maximum of 468 and 62 cfs of runoff are observed in the first and second periods respectively.

.2 Statistical Characteristics of the Data

The histograms of rainfall and runoff sequences are shown in Fig. 2.1. About 80% of rainfall was

TABLE 2.1 DETAILS OF LOCATION OF STATIONS AND DATA USED

STATION	DATA	TYPE OF DATA	DURATION OF DATA	STATION LOCATION AND DETAILS	SOURCE OF DATA
Elgin, near Arlington Heights, Illinois.	Precipitation (inches)	Daily	1952-53 and 1963-64	Lat. 42°02' Long. 88'17'	National Weather Service
Salt Creek, near Arlington Heights, Illinois.	Surface Runoff (cfs)	Daily	1952-53 and 1963-64	Lat. 42°03'02" Long. 88°00'37"	U.S.G.S. Water Supply Papers
Drainage area: 33.7 sq. mi.					

less than 0.1 in. during 1952-53 and about 88% was less than 0.18 in. during 1963-64. About 63% of runoff was less than 10.50 cfs during 1952-53 and about 82% of runoff data was less than 18.02 cfs during 1963-64. The basic statistical characteristics of rainfall and runoff data used is given in Table 2.2. The mean daily rainfall in the two periods is approximately equal whereas the variance of daily rainfall

TABLE 2.2 STATISTICAL CHARACTERISTICS OF THE DATA USED

DATA	DURATION	MEAN	VARIANCE	COEFF. OF SKEWNESS	COEFF. OF KURTOSIS	MAX. VALUE
Rainfall (inches)	1952-54 1963-65	0.098 0.103	0.099 0.069	7.235 4.666	74.919 32.478	4.470 2.780
Runoff (cfs)	1952-54 1963-65	13.641 17.701	1109 1112	6.636 3.392	65.102 17.058	448 262

is smaller during 1963-64 than during 1952-53. The coefficient of skewness and kurtosis of rainfall data are also smaller during 1963-64. The mean runoff shows an increase during 1963-64. The mean runoff shows an increase during 1963-64 compared to the 1952-53 period and the variance is almost the same during the two periods. Similarly the coefficients of skewness and kurtosis are smaller for the runoff data in the period 1963-64, than in the period 1952-53.

The correlation coefficients of rainfall and runoff series and the cross-correlation coefficients of rainfall and runoff yield valuable information about the rainfall-runoff process. The cross-correlation coefficients indicate the inherent memory of the process. These correlation coefficients were computed as follows.

Let us consider a runoff sequence $Z(t)$, $t = 1,2, \ldots N$ where N is the number of observations. Let $\bar{Z}$ by the estimated mean value of the $Z(t)$ series.

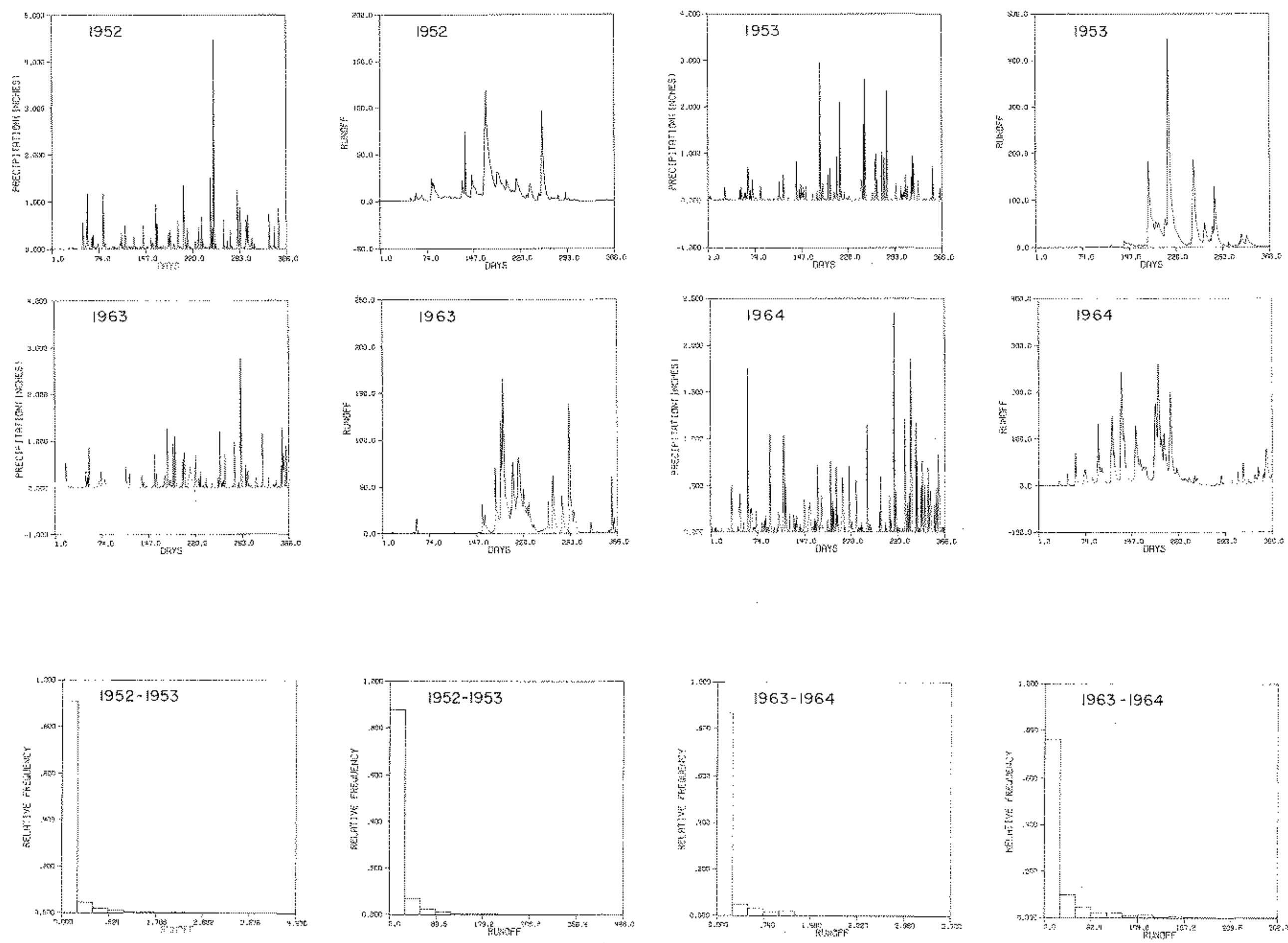


FIG. 2.1 OBSERVED DAILY RAINFALL AND RUNOFF (WATER YEARS) AND HISTOGRAM OF THE OBSERVED RAINFALL AND RUNOFF OF THE SALT CREEK BASIN

$$\bar{Z} = \frac{1}{N}\sum_{t=1}^{N} Z(t) \tag{2.1}$$

The autocovariance $\hat{\gamma}(k)$ of the series $Z(t)$ is given by

$$\hat{\gamma}(k) = \frac{1}{N-k}\sum_{t=1}^{N-k} [Z(t)-\bar{Z}][Z(t+k)-\bar{Z}] \tag{2.2}$$

The autocorrelation coefficient $\hat{r}(k)$ of the runoff sequence is defined by eq. 2.3.

$$\hat{r}(k) = \gamma(k)/\gamma(0) \tag{2.3}$$

Similarly the autocorrelation coefficients of rainfall sequence $I(t)$ can be computed by replacing $Z(t)$ by $I(t)$ in eqs. 2.1 - 2.3. The cross covariance between the rainfall and runoff sequence is given by eq. 2.4.

$$\hat{\gamma}_{12}(k) = \frac{1}{(N-k)}\sum_{t=1}^{N-k+1} [Z(t+k-1)-\bar{Z}][I(t-1)-\bar{I}] \tag{2.4}$$

The cross-correlation coefficients $\hat{r}_{12}(k)$ of the rainfall-runoff process is defined by eq. 2.5.

$$\hat{r}_{12}(k) = \hat{\gamma}_{12}(k)/\hat{\gamma}_{12}(0) \tag{2.5}$$

The correlograms of rainfall and runoff sequences in the two periods are shown in Fig. 2.2. The correlation in the rainfall sequences is smaller for the 1952-64 data. The correlograms of runoff sequences indicate that the runoff sequences are correlated up to a large number of lags in both the periods. The 95% confidence limits (*Anderson, 1942*) are also shown on the correlograms of Fig. 2.2.

The cross-correlation coefficients between the runoff (leading variable) and rainfall sequences are plotted in Fig. 2.2. The cross correlation between runoff and rainfall is high for smaller lags and becomes smaller for increasing lags. Significant cross correlations between rainfall and runoff are found at smaller lags in both the periods. The cross-correlation coefficients vary from 0.341 at lag zero to 0.095 at lag 8 for the 1952-53 data and from 0.330 at lag zero to 0.114 at lag 2 for the 1963-64 data. The significant cross correlations approximately indicate the memory of the rainfall-runoff system. Therefore, the rainfall-runoff process of the Salt Creek basin can be assumed to have a memory of 8 days during 1952-53 and 2 days during 1963-64. Thus the memory of the rainfall-runoff process has decreased with increased urbanization.

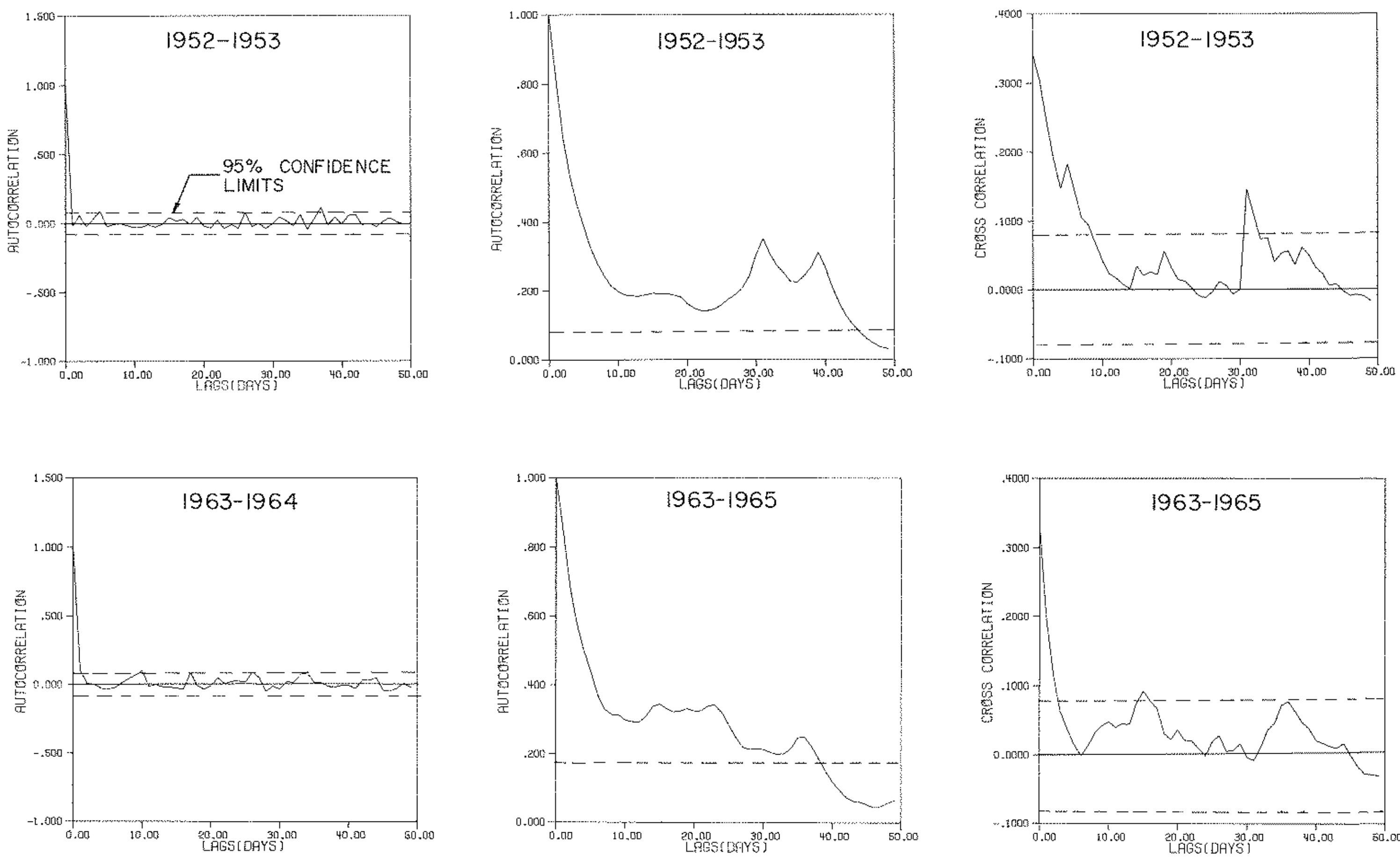


FIG. 2.2 CORRELOGRAMS AND CROSS CORRELATION COEFFICIENTS OF THE OBSERVED RAINFALL AND RUNOFF

CHAPTER III

RECURSIVE PARAMETER ESTIMATION IN THE FUNCTIONAL SERIES MODEL

3.1 Introduction

In modeling the rainfall-runoff process by using a functional series model, several methods based on the multiple regression approach have been used to estimate the kernels. Polynomial approximation of either input or kernels have been used to obtain suitable predictor equations by *Brandstetter and Amorocho (1970)*. The grouped variables of rainfall values, rather than the observed rainfall values have been used in the functional series model by *Bidwell (1971)*. The various parameters of the model are selected by a trial and error procedure.

In this chapter, the functional series model is used to model the rainfall-runoff process. The coefficients in the discrete form of the functional series model are estimated by using a least squares recursive algorithm. The recursive algorithm can also be used to easily update the estimates of the parameters as more input and output information becomes available. The algorithm is particularly useful when analyzing long records of data, such as daily rainfall and runoff series as it requires very little computer storage. The algorithm is also very useful for screening different models. In the estimation scheme discussed below, the data may not be transformed, the data from several stations may be incorporated into the model, and the only parameter that has to be specified is the length of memory of the rainfall-runoff process. Autoregressive terms can also be included into the model. Thus the Volterra Series may be recast as a time series model of rainfall-runoff process.

3.2 The Model and the Estimation Scheme

Let I(k) and Z(k), k=1,2, ... N be the observed daily rainfall and runoff values. The second order nonlinear functional model of the rainfall-runoff process is given in eq. 3.1 for single input and output series,

$$Q(k) = H_0 + \sum_{\tau=0}^{M} H_1(\tau)\, I(k-\tau) + \sum_{\tau_1=0}^{M} \sum_{\tau_2=0}^{M} H_2(\tau_1,\tau_2) I(k-\tau_1)\, I(k-\tau_2) \tag{3.1}$$

where

$Q(k)$ = model output;

M = length of memory.

All the quantities in eq. 3.1 have been defined in Chapter 1. Eq. 3.1 can be rewritten as in eq. 3.2,

$$Q(k) = \underset{\sim}{\alpha}^T(k)\, \underset{\sim}{x}(k), \tag{3.2}$$

where

$$\underset{\sim}{\alpha}^T = \begin{bmatrix} H_0 \\ H_1(0) \\ \cdot \\ \cdot \\ \cdot \\ H_1(M) \\ H_2(0,0) \\ H_2(0,1) \\ H_2(1,1) \\ \cdot \\ \cdot \\ \cdot \\ H_2(M,M) \end{bmatrix} \quad \text{and } \underset{\sim}{x}(k) = \begin{bmatrix} 1 \\ I(k) \\ \cdot \\ \cdot \\ \cdot \\ I(k-M) \\ I(k)\ I(k) \\ I(k)\ I(k-1) \\ \cdot \\ \cdot \\ \cdot \\ I(k-1)\ I(k-1) \\ \cdot \\ \cdot \\ \cdot \\ I(k-M)\ I(k-M) \end{bmatrix} \qquad (3.3)$$

Thus $\underset{\sim}{\alpha}$ is a parameter vector of linear kernels $H_1(\tau)$ and second order kernels $H_2(\tau_1,\tau_2)$. $\underset{\sim}{x}(k)$ is a vector formed by using linear and second order products of rainfall values $I(k)$. In writing eq. 3.2, it is assumed that the second order kernels $H_2(\tau_1,\tau_2)$ are symmetric, i.e.,

$$H_2(\tau_1,\tau_2) = H_2(\tau_2,\tau_1),$$

and this symmetry property reduces the number of second order kernels to be estimated. In a second order model with a memory length of M, the total number n of kernel ordinates $H_1(\cdot)$ and $H_2(\cdot)$ to be estimated in eq. 3.1 or 3.2 is given by

$$n_t = \frac{(M+3)\ (M+2)}{2} \quad . \qquad (3.4)$$

n_t is made up of n_ℓ linear kernels and n_s second order kernels.

$$n_\ell = (M+1) \qquad (3.5)$$

$$n_s = n_t - n_\ell \qquad (3.6)$$

The total number of parameters to be estimated including the constant response term H_0 is

$$n = n_\ell + 1. \qquad (3.7)$$

Thus the $\underset{\sim}{x}$ is $(n\times 1)$ vector and $\underset{\sim}{\alpha}$ is $(1\times n)$ vector.

The second order nonlinear functional model of rainfall-runoff process is given in eq. 3.1. Third or higher order nonlinear models may also be formualted by adding additional rainfall terms into eq. 3.1. Again the symmetry property of the higher order kernels can be taken into account in defining the vectors $\underset{\sim}{\alpha}$ and $\underset{\sim}{x}$.

The model given in eq. 3.2 can also be used for multiple inputs as follows. Let $I_1(k)$ and $I_2(k)$, $k = 1,2, \ldots, N$ be the rainfall values from two stations and let $Z(k)$, $k = 1,2, \ldots, N$ be the runoff value measured at a single station. The model given in eq. 3.2 is still valid for the case of two inputs where the vectors $\underset{\sim}{x}$ and $\underset{\sim}{\alpha}$ are redefined as in eq. 3.8,

$$\underset{\sim}{\alpha} = \begin{bmatrix} \underset{\sim}{\alpha}_1 \\ \\ \underset{\sim}{\alpha}_2 \end{bmatrix} \quad \text{and } x(k) = \begin{bmatrix} x_1(k) \\ \\ x_2(k) \end{bmatrix} \tag{3.8}$$

where $x_i(k)$, $i = 1,2$ are formed by using the i^{th} input and the parameter vector $\underset{\sim}{\alpha}_i$ is defined as in eq. 3.3. The same argument can be extended for more than two inputs. The formulation of the multi-input model given by eq. 3.8, however, does not consider the interaction between the two inputs.

Now assume that the unobservable "true" runoff values $Q(k)$ are corrupted by an additive noise $\eta(k)$, to give the observed runoff values $Z(k)$ as in eq. 3.9.

$$Z(k) = Q(k) + \eta(k) \tag{3.9}$$

From eq. 3.2, we have

$$Q(k) = \underset{\sim}{\alpha}^T \underset{\sim}{x}(k). \tag{3.10}$$

Therefore, eq. 3.10 can be considered as a predictor equation of the runoff process. The noise sequence $\eta(k)$ is assumed to be uncorrelated with $\eta(k-j)$ and $x(k-j)$ $\forall$ $j = 1,2, \ldots$. It is required to estimate the parameter vector $\underset{\sim}{\alpha}$ in eq. 3.10 by using the available rainfall-runoff data. The parameter vector $\underset{\sim}{\alpha}$ can be estimated by minimizing the quadratic functional $J_k(\underset{\sim}{a})$ given in eq. 3.11

$$J_k(\underset{\sim}{a}) = \sum_{j=1}^{k} [Z(j) - \underset{\sim}{a}^T \underset{\sim}{x}(j-1)]^2 + || \underset{\sim}{a} - \underset{\sim}{a}(0) ||^2_{\underset{\sim}{S}^{-1}(0)} \tag{3.11}$$

where

$\underset{\sim}{a}(k)$ = estimate of $\underset{\sim}{\alpha}$ (k)

$\underset{\sim}{S}(k)$ = (n×n) positive definite weighting function matrix

$\underset{\sim}{a}(0)$ = the prior estimate of $\underset{\sim}{\alpha}$ (0)

$\underset{\sim}{S}(0)$ = the prior estimate of $\underset{\sim}{S}(0)$

The parameter vector $\underset{\sim}{\alpha}$ can be estimated by using the recursive algorithm AL.

3.3 Algorithm AL

At any iteration k, $\underset{\sim}{a}(k)$ can be estimated by using the algorithm AL.

$$\underset{\sim}{S}(k+1) = \underset{\sim}{S}(k) - \frac{\underset{\sim}{S}(k)\ \underset{\sim}{x}(k)\ \underset{\sim}{a}^T(k)\ \underset{\sim}{S}(k)}{1 + \underset{\sim}{x}^T(k)\ \underset{\sim}{S}(k)\ \underset{\sim}{x}(k)} \qquad \text{AL}$$

$$\underset{\sim}{a}(k+1) = \underset{\sim}{a}(k) + \underset{\sim}{S}(k+1)\ \underset{\sim}{x}(k)\ [Z(k+1) - \underset{\sim}{a}^T(k)\ \underset{\sim}{x}(k)]$$

The initial estimates of $\underset{\sim}{S}(k)$ and $\underset{\sim}{a}(k)$ can be obtained by using the first k values of Z(k) and I(k). The algorithm can be used for prediction in two ways. The parameter vector $\underset{\sim}{a}(N)$ can be estimated by using N observations. The estimates $\underset{\sim}{a}(N)$ can then be used for prediction by using the rainfall values at N+1, N+2, etc. On the other hand, the parameter vector $\underset{\sim}{a}(N)$ can be updated by using the additional data and these updated parameter estimates can then be used for prediction. Thus the algorithm AL is convenient for real time operation.

The recursive algorithm AL, has been successfully used for estimating the parameters in models of several covariance stationary processes. Successful application of this algorithm in modeling runoff, rainfall and ground water level processes can be found in *Rao and Kashyap* (1973), and *Rao et al.* (1975).

Significance of the parameter estimates:

The parameter vector $\underset{\sim}{a}^N$ is a vector of kernels of various orders and some of these kernels may be insignificant. The significance of the parameter estimates $\underset{\sim}{a}(N)$ can be tested by computing the estimates of the standard errors of parameters. A parameter is considered insignificant if the value of the parameter estimate is less than its standard error. The standard errors of the parameters can be computed by using eq. 3.12,

$$\sigma_a^2 = \underset{\sim}{S}(N)\ \hat{\rho} \tag{3.12}$$

where

$\underset{\sim}{S}(N)$ is obtained from the algorithm AL

N is the number of samples used for estimation of $\underset{\sim}{a}(N)$

ρ equals estimate of the variance of errors $\eta(\cdot)$

The convergence properties of the algorithm are discussed by *Kashyap and Rao* (1976).

CHAPTER IV

A NONLINEAR STOCHASTIC MODEL OF THE RAINFALL-RUNOFF PROCESS

4.1 Introduction

The rainfall and runoff process may be modeled by using the time derivatives of rainfall or runoff. *Kulandaiswamy* (1964) for example, considered a linear differential equation to represent the rainfall-runoff relationship in which time derivatives, up to the third order, of rainfall and runoff were used. The rationale for developing the models of rainfall-runoff process with the time derivatives of rainfall and runoff is simply that the runoff is affected by not only the magnitude of rainfall but also by the rate of occurrence of rainfall. Therefore, in the present study a nonlinear stochastic model of rainfall-runoff process based on the time derivatives of rainfall is formulated. The rainfall and runoff processes are assumed to be random and the model is characterized by the statistical properties of observed rainfall and runoff sequences. Usually time derivatives of rainfall of second or third order are sufficient in the modeling the rainfall-runoff process (*Kulandaiswamy* (1964)). The model and the estimation scheme presented in this chapter are quite different from the models that are available in the literature on rainfall-runoff relationships.

The basic principle of the model is as follows. The model is composed of several subsystems in parallel. The total runoff is a sum of the outputs from these subsystems. Each of these subsystems is nonlinear. The first subsystem utilizes the observed rainfall and runoff values as input and output respectively. This subsystem is called the *zero order* subsystem. Several higher order subsystems act in parallel with the *zero order* subsystem. The input to the second subsystem is the sequence of *first* derivatives of rainfall and the *error* in forecasting the output from the *first* subsystem is considered as the output of this subsystem. Similarly the input to the *third* subsystem is the sequence of *second* derivatives of rainfall and the *error* in forecasting the output from the *second* subsystem is considered as the output of this subsystem. In this way several subsystems are added in parallel. The number of subsystems required to model rainfall-runoff process depends upon the accuracy desired. The error in forecasting the runoff is progressively reduced by each of these subsystems. The schematic representation of the model is shown in Fig. 4.1.

The method of analysis of each of these subsystems is the same except that different inputs and outputs are used in characterization of the subsystems. The optimal output of each of the subsystems is estimated by minimizing the corresponding mean square error and the estimation procedure is based on the statistical theory of filters. The estimation procedure involves the joint probability density functions of rainfall and runoff and their derivatives and these are approximated by orthogonal polynomials. The coefficients in the system equations are computed by using the statistical moments of rainfall and runoff and their derivatives.

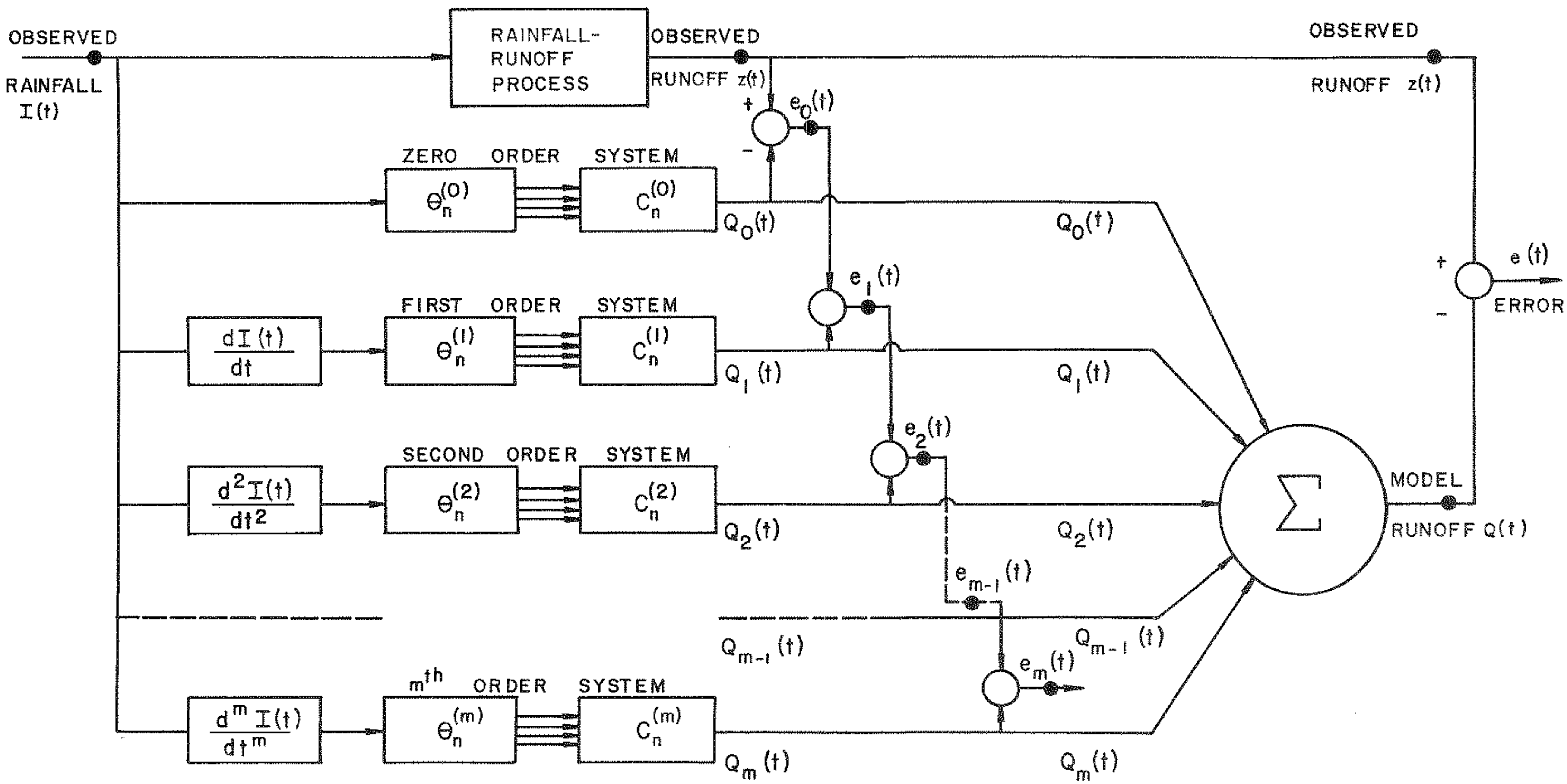


FIG. 4.1 SCHEMATIC REPRESENTATION OF THE NONLINEAR STOCHASTIC MODEL OF RAINFALL-RUNOFF PROCESS

4.2 The Model

Let Z(t) and I(t), t = 1,2, ..., N be the observed runoff and rainfall values. It is assumed that both rainfall and runoff are random functions of time. The relationship between the rainfall and runoff in a single input-single output system may be represented as in eq. 4.1,

$$Q(t) = f\left[I(t),\ I(t),\ I(t),\ \ldots,\ \frac{d^m I(t)}{dt}\right] \tag{4.1}$$

where Q(t) is the model output, m is the highest order of the derivative of rainfall, and f(·) is a nonlinear function. Eq. 4.1 can be rewritten as

$$Q(t) = \sum_{i=0}^{m} f\left(\frac{d^i I(t)}{dt^i}\right) \tag{4.2}$$

(4.3)

$$= \sum_{i=0}^{m} Q_i(t) \tag{4.4}$$

with

$$\frac{d^0 I(t)}{dt^0} = I(t).$$

The total output Q(t) is thus the sum of individual outputs $Q_i(t)$, from each of the systems of various orders, i = 0,1,2, ... m. The *zero order* system is characterized by the observed rainfall and runoff, and the output from the zero order system is $Q_1(t)$ with $e_0(t) = Z(t) - Q_1(t)$. The *first order* system is characterized by the first derivative of rainfall dI/dt and error $e_0(t)$ in forecasting the runoff from the zero order system as input and output respectively. The forecast error from the first order system is designated $e_1(t)$. Similarly the *second order* system is characterized by the second derivative of rainfall and error $e_1(t)$ from the *first order* system as input and output respectively. Several subsystems can be further added with increasing order of time derivative of rainfall as input and the error from the previous subsystem as output. The model output is computed as the sum of outputs from each of these subsystems. Each of these subsystems are nonlinear and are realized as zero memory nonlinear systems. Some preliminary aspects of the zero memory of nonlinear systems have been discussed in Chapter 1. The computation of optimum output of a zero memory system is discussed in the following section. The computation of outputs $Q_0(t)$, $Q_1(t)$, etc., of the subsystems of the nonlinear model of rainfall-runoff process is also discussed in the next section.

4.3 Optimum Output of a Zero Memory System

The optimum output Q(t) of a zero memory nonlinear system is obtained by minimizing the mean square error between the model output Q(t) and the observed output Z(t) by using statistical theory of filters as discussed by *Hubbell (1967)*. The optimum value of Q(t) is equal to the conditional mean of $Z(t)$ given $I(t)$ and is given in eq. 4.5

$$Q(t) = \frac{1}{P_I(i)} \int Z(t)\, P_{IZ}(i,z)\, dz \tag{4.5}$$

where $P_I(i)$ is the probability density function of rainfall and $P_{IZ}(i,z)$ is the joint probability density of the rainfall and runoff.

The model of the rainfall-runoff process given in eq. 4.5 is characterized by the probability density function of rainfall and the joint probability density function of rainfall and runoff. The optimal output of the model can be computed with the knowledge of these two density functions. The joint probability density function $P_{IZ}(i,z)$ is approximated by an orthogonal polynomial and the optimum $Q(t)$ value is evaluated in terms of this orthogonal polynomial.

Let $P_{IZ}(i,Z)$ be approximated by an expansion given in eq. 4.6,

$$P_{IZ}(i,z) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} A_{nm}\, \theta_n(I(t))\xi_m(Z(t))P_I(i)P_Z(z) \tag{4.6}$$

where $\theta_n(I(t))$ and $\xi_m(Z(t))$ are two sets of polynomials which are orthonormal with respect to $P_I(i)$ and $P_Z(z)$, as weighting functions over the whole range of $I(t)$ and $Z(t)$ respectively; $P_I(i)$ and $P_Z(z)$ are the probability density functions of $I(t)$ and $Z(t)$ respectively. Substituting eq. 4.6 in eq. 4.5 we get

$$Q(t) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} A_{nm}\, \theta_n(I(t)) \int Z(t)\, \xi_m(Z(t))P_Z(z)\, dz \tag{4.7}$$

It can be shown that (*Hubbell, (1967)*),

$$\int Z(t)\xi_m(Z(t))\, P_Z(z)\, dz \begin{cases} = \bar{Z} & \text{for } m = 0 \\ = \sigma_Z & \text{for } m = 1 \\ = 0 & \text{for } m \neq 0,\ m \neq 1 \end{cases} \tag{4.8}$$

and

$A_{00} = \quad = 1$

$A_{0n} \quad = A_{n0} = 0 \qquad \text{for } n \neq 0$

$\theta_0(I(t)) = 1.$

In eq. 4.8, $\bar{Z}$ and σ_Z^2 are respectively the mean and variance of $Z(t)$. Now using eq. 4.8 in eq. 4.7 we can show that

$$Q(t) = \sum_{n=1}^{\infty} A_{n1}\, \theta_n(I(t))\, \sigma_Z + \bar{Z} \tag{4.8a}$$

Truncating the series in eq. 4.8a at the Rth term, we have

$$Q(t) = \sum_{n=1}^{R} A_{n1}\ \theta_n(I(t))\ \sigma_Z + \bar{Z} \qquad (4.9)$$

Equation 4.9 can be rewritten as eq. 4.10,

$$Q(t) = \sum_{n=0}^{R} C_n\ \theta_n(I(t)) \qquad (4.10)$$

where $C_0 = \bar{Z}$ and $C_n = A_{n1}\ \sigma_Z$, $n = 1,2, \ldots, R$

The orthogonal polynomials θ_n are obtained by using eq. 4.11,

$$\begin{aligned} \theta_n(I(t)) &= a_{nn}I^n(t) - \sum_{r=0}^{n-1} a_{nr}\ \theta_r(I(t)); \\ \theta_0(I(t)) &= 1, \\ \theta_1(I(t)) &= \frac{I-\bar{I}}{\sigma_I} \end{aligned} \qquad (4.11)$$

where σ_I^2 is the variance of I(t) and $\bar{I}$ is the mean value of I(t). These polynomials can be generated by using the following recursive relationships,

$$\theta_n(I(t)) = a_{nn}\left\{I(t)^n - \sum_{n=0}^{n-1} \overline{[I(t)^n\ \theta_r(I(t))]}\theta_r(I(t))\right\} \qquad (4.12)$$

and

$$a_{nn} = \left\{I(t)^{2n} - \sum_{r=0}^{n-1} \overline{[I(t)^n\ \theta_r(I(t))]^2}\right\}^{-\frac{1}{2}} .$$

Similarly the coefficients C_n can be computed by using the relationship given in eq. 4.13,

$$C_n = a_{nn}\ \overline{I(t)^n Z(t)} - \sum_{r=0}^{n-1} \overline{[I(t))^n\ \theta_r(I(t))]}A_{r1}\ \sigma_Z; \qquad (4.13)$$

$$n = 1,2, \ldots R$$

and

$$\sigma_Z\ A_{n1} = a_{nn}\ \overline{I(t)^n Z(t)} - \sum_{r=1}^{n-1} a_{nr}\ \overline{\theta_r(I(t))Z(t)} .$$

The coefficients θ_n (I(t)) and the polynomials C_n can be calculated by using eqs. 4.12 and 4.13. The system configuration is shown in Fig. 4.2.

Now returning to the nonlinear model given in eq. 4.1 or 4.2, the outputs $Q_0(t)$, $Q_1(t)$, $Q_2(t)$, etc., can be computed for each of the subsystems as follows.

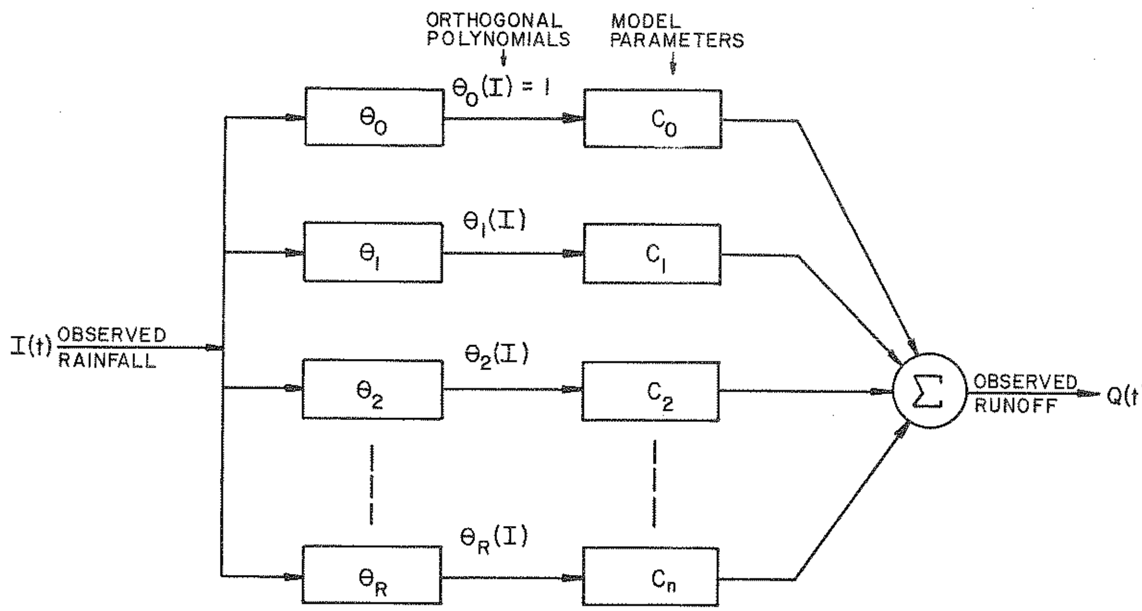


FIG. 4.2 REALIZATION OF A ZERO MEMORY NONLINEAR SYSTEM

Zero order system:

The zero order system is characterized by the observed rainfall and runoff with coefficients $C_n^{(0)}$, $n = 0,1,2, \ldots R$. The output and error of a zero order system are respectively designated by $Q_0(t)$ and $e_0(t)$

$$\left.\begin{aligned} Q_0(t) &= \sum_{n=0}^{R} C_n^{(0)} \; \theta_n^{(0)} \; (I(t)) \\ e_0(t) &= Z(t) - Q_0(t) \end{aligned}\right\} \qquad (4.14)$$

The coefficients $C_n^{(0)}$ and $\theta_n^{(0)}$ $(I(t))$ are computed by using eqs. 4.12 and 4.13.

First order system:

The first order system is characterized by the first time derivative of rainfall as input and the error $e_0(t)$ from the zero order system as output. The output and error of a first order system are respectively designated by $Q_1(t)$ and $e_1(t)$.

$$\left.\begin{aligned} Q_1(t) &= \sum_{n=0}^{R} C_n^{(1)} \; \theta_n^{(1)} \left(\frac{dI(t)}{dt}\right) \\ e_1(t) &= e_0(t) - Q_1(t) \end{aligned}\right\} \qquad (4.15)$$

The coefficients $C_1^{(1)}$ and polynomials $\theta_n^{(1)} \left(\frac{dI(t)}{dt}\right)$ are computed by using eqs. 4.12 and 4.13 with $\frac{dI(t)}{dt}$ and $e_0(t)$ sequences replacing $I(t)$ and $Z(t)$ respectively.

Second order system:

The second order system is characterized by the second time derivative of rainfall as input and the error $e_1(t)$ from the first order system as output. The output and the error from the second order system are respectively designated by $Q_2(t)$ and $e_2(t)$.

$$\left.\begin{aligned} Q_2(t) &= \sum_{n=0}^{R} C_n^{(2)} \; \theta_n^{(2)} \left(\frac{d^2I(t)}{dt^2}\right) \\ e_2(t) &= e_1(t) - Q_2(t) \end{aligned}\right\} \qquad (4.16)$$

The coefficients $C_n^{(2)}$ and $\theta_n^{(2)} \left(\frac{d^2I(t)}{dt^2}\right)$ are computed by using eqs. 4.12 and 4.13 with $\frac{d^2I(t)}{dt^2}$ and $e_1(t)$ sequences replacing $I(t)$ and $Z(t)$ respectively.

The above process may be continued for orders $m > 2$ also. For an m^{th} order system, the m^{th} time derivative of rainfall is used as input and the error $e_{m-1}(t)$ from the $(m-1)^{st}$ order is used as output. The output of an m^{th} order system is given in eq. 4.17.

$$Q_m(t) = \sum_{n=0}^{R} C_n^{(m)} \theta_n^{(m)} \left(\frac{d^m I(t)}{dt^m}\right)$$
$$e_m(t) = e_{m-1}(t) - Q_m(t) \tag{4.17}$$

The coefficients $C_n^{(m)}$ and polynomials $\theta_n^{(m)}$ $(\cdot)$ are computed by using eqs. 4.12 and 4.13 with m^{th} derivative of rainfall replacing I(t) and $e_{m-1}(t)$ replacing Z(t). The total output of all the subsystems is now given in eq. 4.18.

$$Q(t) = \sum_{i=1}^{m} Q_i(t)$$

or

$$Q(t) = \sum_{i=1}^{m} \sum_{n=0}^{R} C_n^{(i)} \theta_n^{(i)} \left(\frac{d^i I(t)}{dt^i}\right) \tag{4.18}$$

The model given by eq. 4.18 is schematically shown in Fig. 4.2.

Since the computation of time derivative of a sequence I(t) at any particular instant t involves previous I(t) values, it can be argued that the output given by eq. 4.18 depends upon the present and previous values of input. In a model such as that in eq. 4.18, several past values of inputs are used in the computation of various derivatives and hence the runoff from the model is computed by using several past values of inputs. This aspect is not explicitly seen in eq. 4.18.

It was noted in the previous section that each of the subsystems with output $Q_0(t)$, $Q_1(t)$, etc., are nonlinear because the orthogonal polynomials $\theta_n(I(t))$, n = 0,1,2, ..., R are power series in I(t) up to R^{th} power as computed by using eq. 4.12. The nonlinear model is thus realized as a parallel combination of several subsystems and each of these subsystems is a zero memory nonlinear system. Both the observed rainfall and rainfall time derivatives are used as inputs. The value of m, which is the highest order of the rainfall derivative, depends upon the accuracy desired.

Each of the subsystems is characterized by coefficients $C_n^{(i)}$, i = 0,1,2, ... M and n = 0,1,2, ..., R where R is the finite number of terms in eq. 4.10. Hence in any analysis, the values of R and m have to be specified.

The derivatives of rainfall values were computed by using the following formulae (*Salvadori and Baron (1961)*). Let I(t), t = 1,2, ..., N be the sequence of rainfall observed at intervals Δt. In the present case Δt is unity.

Forward difference formula: $\left.\frac{dI(t)}{dt}\right|_{t=1} = \frac{1}{2\Delta t}[-3I(1) + 4I(2) - I(3)]$

Central difference formula: $\frac{dI(t)}{dt} = \frac{1}{2\Delta T}[I(t) - I(t-1)];\ 1 = 2,3, \ldots, N-1$

Backward difference formula: $\left.\frac{dI(t)}{dt}\right|_{t=N} = \frac{1}{2\Delta T}[I(N-2) - 4I(N-1) + 3I(N)]$

The higher order time derivatives of rainfall were similarly computed.

CHAPTER V

ANALYSIS OF THE EFFECTS OF URBANIZATION ON RUNOFF CHARACTERISTICS

5.1 Introduction

Urbanization changes the runoff characteristics significantly. The effects of urbanization on runoff have been investigated by various methods as discussed in Chapter I. As discussed earlier, most of the previous studies are based on analysis of storms. In this chapter, rainfall-runoff models which are discussed in Chapters 3 and 4 will be used to investigate the effects of urbanization on daily runoff.

Rainfall-runoff data from Salt Creek watershed from two different periods, one corresponding to a less urbanized (water years 1952-53) and the other corresponding to a heavily urbanized period (water years 1963-64) are used in the present study. These two periods will be respectively called the first and second periods. The following procedure is used to evaluate the effect of urbanization on daily runoff characteristics.

(a) As urbanization changes the runoff characteristics, the model fitted to the data during the first period cannot be expected to perform well as a runoff predictor for the second period and vice versa. This hypothesis will be tested by separately fitting valid models by using the data from the two periods and evaluating the prediction performances of these models in the other period. The valid models for the rainfall-runoff data will be obtained as discussed in Chapters III and IV.

(b) The effects of urbanization on runoff characteristics may be investigated by simulation studies. For example, the valid models of the rainfall-runoff process developed by using the data from a period such as the first period can be used to generate synthetic runoff in the second period by using the rainfall in the second period and the parameters estimated by first period data. The simulated runoff sequence can be used to evaluate the difference in runoff characteristics in the two periods brought about by urbanization. The valid models, however, should be tested for good simulation performance. The effect of urbanization can be quantified by comparing several properties of the simulated runoff series in the two periods, such as mass curves, flow duration curves, frequency analysis of exceedence values, and flood peaks etc.

The basic characteristics of the Salt Creek basin data have been discussed in Chapter II. The runoff characteristics such as the peak runoff, mean daily runoff and skewness coefficient are different in the two periods. The cross correlation coefficients of the rainfall-runoff are significant upto lags of 7 and 2 days for the data during 1952-53 and 1963-64 respectively. Thus the memory of the rainfall-runoff process is reduced from about 7 days to 2 days, which may be attributed to the effects of urbanization of the basin.

5.2 Analysis of the Effects of Urbanization on Runoff Prediction-Functional Series Model

The functional series model, discussed in Chapter III was used to fit models to the observed rainfall-runoff data from the Salt Creek basin during the first and second periods. Separate valid models were obtained for these two periods by using the measured rainfall-runoff values in the two periods. These models were used for prediction of runoff by using the observed rainfall from the other period. For example the kernel functions estimated by using the rainfall-runoff data from the first period are used with the observed rainfall in the second period to predict the runoff in the second period and vice versa. The prediction performance of the models were evaluated. The validity of the models was ascertained by an analysis of regeneration results.

The regeneration and prediction performance of the models were evaluated as follows. Let the error in regeneration or prediction be e(t), as defined in eq. 5.1 where Z(t) and Q(t) are respectively the observed and regenerated or predicted runoff.

$$e(t) = Z(t) - Q(t) \tag{5.1}$$

The performance of the model can be evaluated by considering the statistical properties of the e(t), or by other measures. In the present study, the statistics R_1, R_2 and R_3 and the plots of observed and regenerated (or predicted) runoff values are used to evaluate the performance of the model. The statistics R_1, R_2 and R_3 are defined below.

$$R_1 = \frac{\text{mean error}}{\text{mean observed runoff}} = \frac{\sum_{t=1}^{N} e(t)/N}{\sum_{t=1}^{N} Z(t)/N} \tag{5.2}$$

$$R_2 = \frac{\text{error mean square}}{\text{runoff mean square}} = \frac{\sum_{t=1}^{N} e^2(t)/N}{\sum_{t=1}^{N} Z^2(t)/N} \tag{5.3}$$

R_3 = correlation coefficient between observed and computed runoff.

$$R_3 = \frac{N \sum_{t=1}^{N} Z(t)\, Q(t) - \bar{Z}\, \bar{Q}\, N^2}{\sum_{t=1}^{N} (Z(t)-\bar{Q})^2 \sum_{t=1}^{N} (Z(t)-\bar{Q})^2} \tag{5.4}$$

In eqs. 5.2-5.4 $\bar{Z}$ and $\bar{Q}$ are respectively the mean observed and computed runoff. If the model gives good regeneration or prediction performance, the quantities R_1 and R_2 should be small compared to unity and R_3 should be close to unity. The time to peak and peak runoff of the computed and observed runoff should also be close to each other.

Functional series models with first and second order terms and with only first order terms were used in the analysis. These two models will be respectively called nonlinear and linear models. The model is given in eq. 3.1 and the recursive algorithm AL discussed in Chapter III was used to estimate the kernels $H_1(\tau)$ in the linear model, and $H_1(\tau)$ and $H_2(\tau_1,\tau_2)$ in the nonlinear model. The second order kernels are assumed to be symmetric. The bias term H_0 in eq. 3.1 was dropped as it was not significant.

In the linear and nonlinear models the length of memory (M) has to be specified. The length of memory was varied from 2 to 15 days, and the ratio R_2 of the error mean square to the runoff mean square was computed in linear and nonlinear models for the different M values. The value of R_2 for different memory lengths are plotted in Fig. 5.1. From Fig. 5.1 it is apparent that 7 and 2 days of memory result in minimum values of R_2 for the two periods. The cross correlation coefficients between rainfall-runoff indicate approximately the same results. The functional series models (linear and nonlinear) for the two periods can now be written as in eq. 5.5.

First period

Linear model (L1):

$$Q(t) = \sum_{\tau=0}^{7} H_1(\tau)I(t-\tau)$$

Nonlinear model (N1):

$$Q(t)=\sum_{\tau=0}^{7} H_1(\tau)I(t-\tau) + \sum_{\tau_1=0}^{7} \sum_{\tau_2=\tau_0}^{7} H_2(\tau_1,\tau_2)I(t-\tau_1)I(t-\tau_2)$$

Second period (5.5)

Linear model (L2):

$$Q(t) = \sum_{\tau=0}^{2} H_1(\tau)I(t-\tau)$$

Nonlinear model (N2):

$$Q(t) = \sum_{\tau=0}^{2} H_1(\tau)I(t-\tau) + \sum_{\tau_1=0}^{2} \sum_{\tau_2=\tau_0}^{2} H_2(\tau_1,\tau_2)I(t-\tau_1)I(t-\tau_2)$$

The parameters of the above models of the rainfall-runoff process were separately estimated for the 1952-53 and 1963-64 data by using the recursive algorithm AL discussed in Chapter III. The parameter estimates in both linear and nonlinear models are listed in Table 5.1 for the models given in eq. 5.5. The standard errors of the parameter estimates are also given in Table 5.1. The insignificant parameters are underscored in Table 5.1. All the first order kernels are significant in both linear

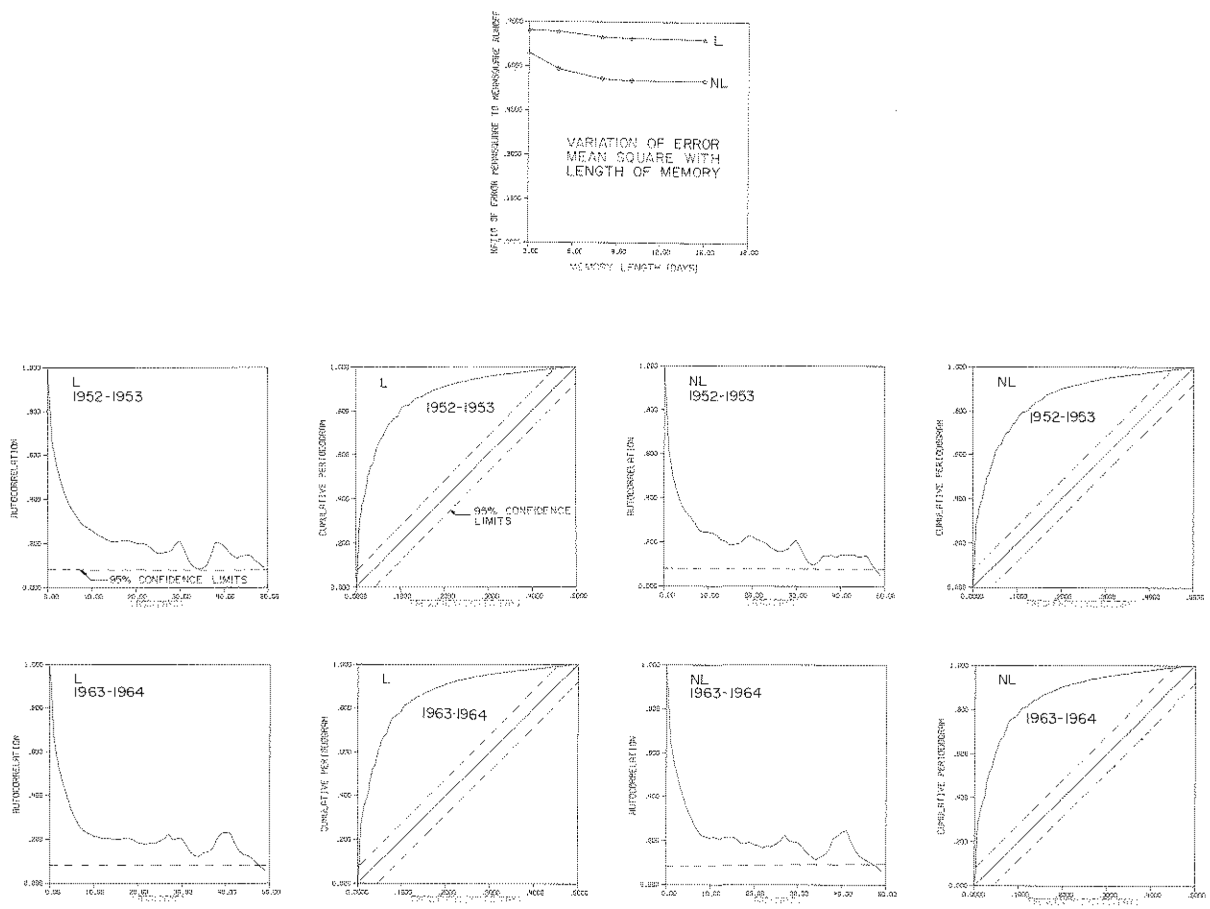


FIG. 5.1 VARIATION OF R_2 WITH THE LENGTH OF MEMORY, CORRELOGRAMS AND CUMULATIVE PERIODOGRAMS OF REGENERATION ERRORS FROM THE LINEAR AND NONLINEAR FUNCTIONAL SERIES MODELS

TABLE 5.1a PARAMETER ESTIMATES IN FUNCTIONAL SERIES MODELS (1952-53 DATA)

Linear Model

	$\tau \rightarrow$	0	1	2	3	4	5	6	7
	$H_1(\tau)$	34.30 (3.70)	30.77 (3.70)	22.48 (3.72)	18.09 (3.68)	12.92 (3.68)	13.80 (3.72)	11.66 (3.72)	8.54 (3.70)

Nonlinear Model

	$\tau \rightarrow$	0	1	2	3	4	5	6	7
	$H_1(\tau)$	33.15 (28.24)	42.43 (29.65)	42.00 (28.94)	40.12 (26.78)	24.55 (26.47)	30.27 (27.37)	15.44 (27.30)	4.26 (27.83)

Kernels	$\tau_2 \rightarrow$	0	1	2	3	4	5	6	7
$H_2(\tau_1,\tau_2)$	$\tau_1=0$	-59.85 (15.11)	-93.05 (79.65)	-5 .27 (38.41)	-105.62 (83.11)	11.72 (36.14)	19.25 (33.87)	94.53 (110.52)	19.87 (96.72)
	1		-1.99 (15.32)	-85.78 (79.75)	-4.88 (51.79)	-95.61 (81.49)	-5.20 (36.79)	10.47 (33.87)	91.89 (110.52)
	2			-4.79 (15.07)	-49.22 (80.24)	16.05 (27.29)	-209.35 (121.97)	-52.40 (36.46)	5.05 (34.02)
	3				-1.95 (7.99)	-49.05 (79.69)	-10.79 (23.63)	-148.52 (121.31)	-4.76 (35.47)
	4					-0.16 (8.028)	-17.08 (84.90)	-5.83 (23.26)	-104.97 (121.55)
	5						1.021 (8.79)	-36.67 (9.22)	-11.41 (23.62)
	6							1.34 (8.64)	-20.81 (91.23)
	7								2.92 (8.49)

TABLE 5.1b PARAMETER ESTIMATES IN FUNCTIONAL SERIES MODELS (1963-64 DATA)

Linear Model

	$\tau \rightarrow$	0	1	2
	$H_1(\ )$	50.975 (4.80)	28.99 (4.88)	22.77 (4.80)

Nonlinear Model

	$\tau \rightarrow$	0	1	2
	$H_1(\ )$	88.94 (11.66)	56.12 (11.22)	55.48 (10.86)

Kernels	$\tau_2 \rightarrow$	0	1	2
$H_2(\tau_1,\tau_2)$	$\tau_1=0$	-24.09 (6.67)	-3.69 (16.65)	-65.22 (23.92)
	1		22.74 (7.47)	3.78 (16.73)
	2			-23.98 (7.33)

and nonlinear models for the first period whereas several second order kernels $H_2(\tau_1,\tau_2)$ are insignificant. All the kernels except $H_2(1,2)$ and $H_2(0,1)$ are significant in both the linear and nonlinear models for the second period.

The estimated kernels and the corresponding rainfall values were used to regenerate runoff. The rainfall values which were used in the estimation of kernels are used with the resulting kernels to compute the runoff Q_r. The regeneration residuals $e_r(t)$ were computed by eq. 5.6,

$$e_r(t) = Z(t) - Q_r(t) \tag{5.6}$$

where $Z(t)$ is the observed runoff and $Q_r(t)$ is the model output. The errors $e_r(t)$ were tested for whiteness. The whiteness tests are discussed in *Bax and Jenkins (1970)*. Presently, we will discuss only the results. The correlogram and the cumulative periodograms of the errors are shown in Fig. 5.1. The cumulative periodograms indicate that the $e_r(t)$ sequences are not free of periodicies. The results of *Portmanteau* test on errors are summarized in Table 5.2. The results from both of these tests indicate that the regeneration errors are correlated. The statistics R_1-R_3 of linear and nonlinear model errors for the two periods are summarized in Table 5.3. These statistics indicate that the regeneration performance of the nonlinear model is slightly inferior to that of the linear model if R_2 is considered.

As the residual sequence $e_r(t)$ is correlated, the models as given in eq. 5.5 cannot be considered as valid models of the runoff process. Consequently, autoregressive models were fitted to the residual sequences. Several residual models such as those shown in eq. 5.7 were fitted to the residual series.

$$e_r(t) = \sum_{i=1}^{p} \beta_i \; e_r(t-i) + W(t) \tag{5.7}$$

Several different values of p were tried and the corresponding β_i estimates were determined by using the algorithm AL. A fifth order AR model was found adequate to characterize the residual series. The parameter estimates β_i and their standard errors are given in Table 5.4. The insignificant parameter estimates are underscored in Table 5.4. Most significantly, the parameter β_1 is larger and significant in both the (L1 and N1) models for the residuals from the first period data, and the parameters β_1 and β_2 are significant in the models (L2 and N2) for the residuals from the second period data.

The error models (eq. 5.7) were incorporated into the functional series models given in eq. 5.5 and the resulting models are called *overall* models. The runoff estimated by using the overall model is designated $\hat{Q}(t)$. The overall models are given in eq. 5.8.

TABLE 5.2 RESULTS OF *PORTMANTEAU* TEST ON REGENERATION ERRORS

DATA WATER YEARS	MODEL	LAG	TEST STATISTICS	CRITICAL VALUE AT α = 0.05	DECISION
1952-53	L1	5	783.44	11.07	R*
	N1	5	775.57	11.07	R
1963-64	L2	5	787.52	11.07	R
	N2	5	912.04	11.07	R

*Reject the hypothesis that $e_r(t)$ are uncorrelated.

TABLE 5.3 REGENERATION AND PREDICTION RESULTS OF LINEAR AND NONLINEAR MODELS

DATA WATER YEARS	MODEL	REGENERATION* OR PREDICTION	MEAN RUNOFF (cfs)	MEAN SQUARE RUNOFF $\times 10^3 (cfs)^2$	ERROR MEAN (cfs)	ERROR MEAN SQUARE $\times 10^3 (cfs)^4$	R_1	R_2	R_3
1952-53	L1	R(1952-53)	13.640	1.294	-1.317	0.720	-0.097	0.556	0.589
		P(1963-64)	17.700	1.423	1.031	0.009	0.109	0.709	0.342
	N1	R(1952-53)	13.640	1.294	-2.389	0.588	-0.175	0.454	0.685
		P(1963-64)	17.700	1.423	5.775	1.672	0.326	1.174	0.205
1963-64	L2	R(1963-64)	17.700	1.294	7.098	1.028	0.401	0.722	0.370
		P(1952-53)	13.640	1.423	3.538	0.856	0.259	0.662	0.507
	N2	R(1963-64)	17.700	1.294	3.402	0.945	0.192	0.663	0.425
		P(1952-53)	13.640	1.423	2.643	1.701	0.194	1.315	0.002

*R = Regeneration. P = Prediction.

TABLE 5.4 PARAMETER ESTIMATES OF THE RESIDUAL MODELS

DATA WATER YEARS	MODEL	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$
1952-53	L1	0.858	-0.095	-0.033	0.027	0.014
		(0.046)*	(0.060)	(0.060)	(0.061)	(0.046)
	N1	0.816	-0.042	-0.038	0.053	0.022
		(0.044)	(0.057)	(0.057)	(0.057)	(0.044)
1963-64	L2	0.908	-0.110	-0.031	0.055	0.025
		(0.043)	(0.059)	(0.059)	(0.059)	(0.043)
	N2	0.883	-0.096	-0.011	0.059	0.019
		(0.043)	(0.057)	(0.057)	(0.057)	(0.043)

*Standard errors of the estimates are shown in parentheses.

First period:

Linear model (ML1):

$$\hat{Q}(t) = \sum_{\tau_1=0}^{7} H_1(\tau)I(t-\tau_1) + \sum_{i=1}^{5} \beta_i\, e_r(t-i)$$

Nonlinear model (MN1):

$$\hat{Q}(t) = \sum_{\tau_1=0}^{7} H_1(\tau)I(t-\tau_1) + \sum_{\tau_1=0}^{7}\sum_{\tau_2=0}^{7} H_2(\tau_1,\tau_2)I(t-\tau_1)I(t-\tau_2) + \sum_{i=1}^{5}\beta_i\, e_r(t-i)$$

Second period:

Linear model (ML2):

$$\hat{Q}(t) = \sum_{\tau_1=0}^{2} H_1(\tau)I(t-\tau_1) + \sum_{i=1}^{5} \beta_i\, e_r(t-i)$$

Nonlinear model (MS2):

$$\hat{Q}(t) = \sum_{\tau_1=0}^{2} H_1(\tau)I(t-\tau_1) + \sum_{\tau_1=0}^{2}\sum_{\tau_2=0}^{2} H_2(\tau_1,\tau_2)I(t-\tau_1)I(t-\tau_2) + \sum_{i=1}^{5}\beta_i\, e_r(t-i) \qquad (5.8)$$

The correlograms and the cumulative periodograms of the residuals W(t) (eq. 5.7) shown in Fig. 5.2 demonstrate that the W(t) sequence may be considered uncorrelated and without periodicities. The results of *Portmanteau* test on W(t) given in Table 5.5 also indicate that the residuals W(t) can be considered to be uncorrelated up to 50 lags. The histograms of the residuals W(t) shown in Fig. 5.2 are such that they can be considered to be approximately normally distributed with zero mean. The models for the first period given in eq. 5.8 were used for predicting runoff in the second period by using the rainfall values of the second period and vice versa. The regeneration and prediction performance of the overall models (eq. 5.8) were evaluated and the results are summarized in Table 5.6.

A few remarks about the difference between regeneration and prediction are in order. Let us consider the overall model ML1.

Let the estimates of kernels $H_1(\tau)$, $\tau = 0,1, \ldots, 7$, and the β_i coefficients be obtained by using the 1952-53 rainfall and runoff data. The rainfall values of the 1952-53 period may be used, along with $H_1(\tau)$ and β_i estimated by using the 1952-53 data to regenerate the runoff during 1952-53. Let the regenerated runoff be designated $\hat{Q}_r(t)$ and the residuals be $e_r(t)$. Let Z(t) be the observed runoff. Then,

$$e_r(t) = Z(t) - \hat{Q}_r(t) = W(t),$$

$$= Z(t) - \sum_{\tau=0}^{7} H_1(\tau)I(t-\tau) + \sum_{i=1}^{5}\beta_i e_r(t-i). \qquad (5.9)$$

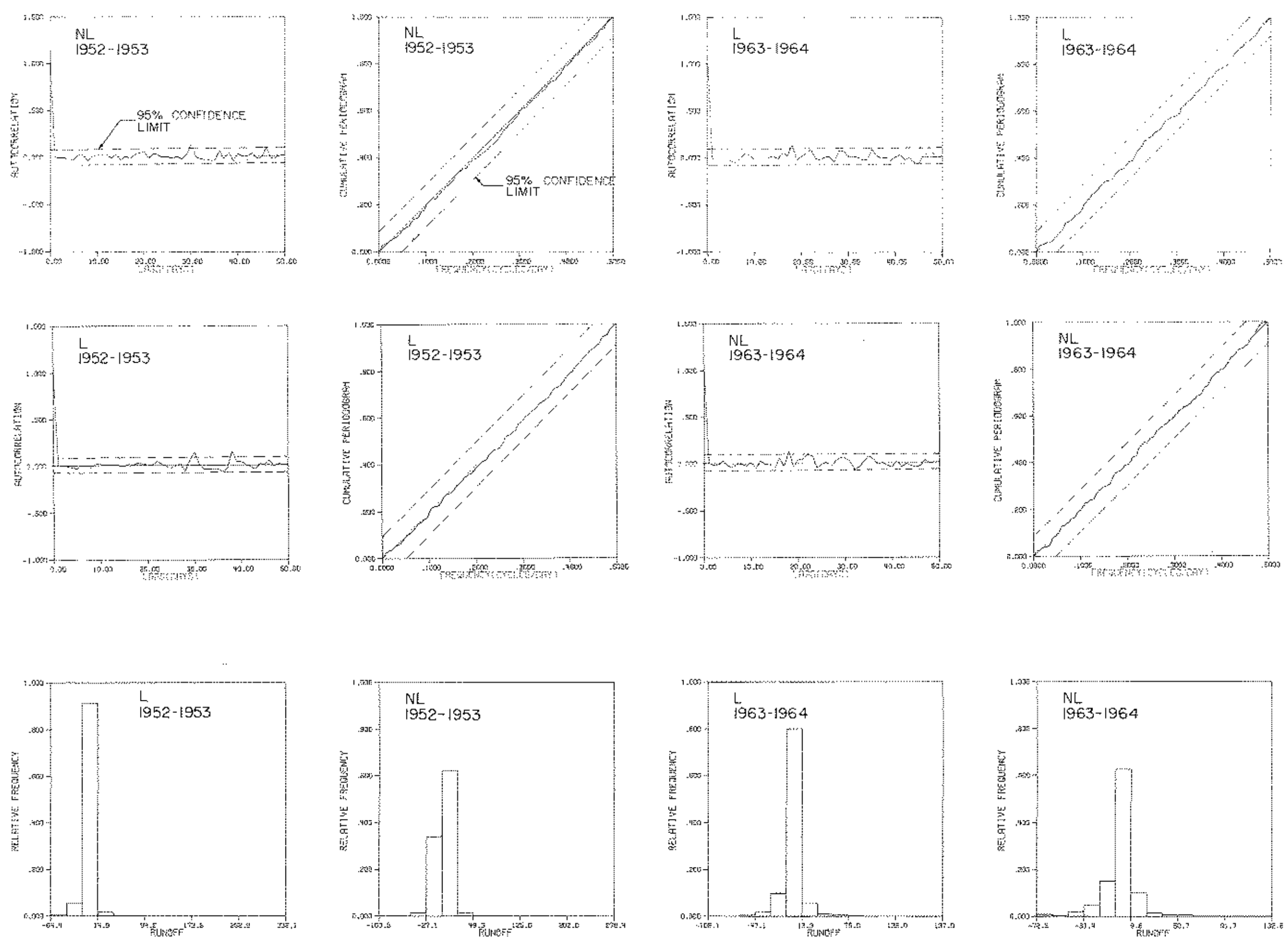


FIG. 5.2 CORRELOGRAMS, CUMULATIVE PERIODOGRAMS AND HISTOGRAMS OF REGENERATION ERRORS FROM THE OVERALL LINEAR AND NONLINEAR FUNCTIONAL SERIES MODELS

TABLE 5.5 RESULTS OF *PORTMANTEAU* TEST ON W(t) SEQUENCE

DATA WATER YEARS	MODEL	LAG (DAYS)	TEST STATISTIC	CRITICAL VALUE $\alpha = 0.05$	DECISION
1952-53	ML1	50	62.40	67.50	A*
	MN1	50	49.05	67.50	A
1963-64	ML2	50	60.35	67.50	A
	MN2	50	67.70	67.50	A

*A = accept the hypothesis that W(t) sequence is uncorrelated.

TABLE 5.6 REGENERATION AND PREDICTION RESULTS OF MODELS ML1, MN1, ETC.

WATER YEARS	MODELS	REGENERATION(R) OR PREDICTION(P)	MEAN RUNOFF (cfs)	MEAN SQUARE RUNOFF $(cfs)^2 \times 10^{-3}$	ERROR MEAN (cfs)	ERROR MEAN SQUARE $\times 10^3 (cfs)^2$	R_1	R_2	R_3
1952-53	ML1	R(1952-53)	13.64	1.294	-0.222	0.238	-0.016	0.184	0.890
		P(1963-64)	17.70	1.423	0.328	0.240	0.018	0.169	0.904
	MN1	R(1952-53)	13.64	1.294	-0.443	0.220	-0.032	0.170	0.904
		P(1963-64)	17.70	1.423	1.096	0.546	0.062	0.383	0.822
1963-64	ML2	R(1963-64)	17.70	1.294	0.930	0.266	0.052	0.187	0.877
		P(1952-53)	13.64	1.423	0.544	0.323	0.040	0.250	0.877
	MN2	R(1963-64)	17.70	1.294	0.494	0.280	0.028	0.197	0.892
		P(1952-53)	13.64	1.423	0.385	0.771	0.028	0.596	0.767

The residuals $e_r(t)$ in eqs. 5.9 and 5.6 are different as they arise from different (L1, ML1; L2, ML2 etc.) models.

The estimates $H_1(\tau)$ and β_i of ML1 estimated by using 1952-53 data may be used to predict runoff in the second period (1963-64) as the rainfall and runoff values in the second period are known. Let the predicted runoff be designated $\hat{Q}_p(t)$ and the prediction error be $e_p(t)$, as in eq. 5.10. The first few (p in eq. 5.7) e_p values were assumed to be zero. The prediction error sequence $e_p(t)$

$$e_p(t) = Z(t) - \hat{Q}_p(t)$$

$$= Z(t) - \sum_{\tau=0}^{7} H_1(\tau) I(t-\tau) + \sum_{i=1}^{5} \beta_i \; e_p(t-i) \qquad (5.10)$$

will have characteristics similar to the residual sequence $e_r(t)$ if the process Z(t) were to remain approximately the same during the two periods. Otherwise they would be different. Consequently, a comparison of the characteristics of the $e_r(t)$ and $e_p(t)$ sequences would provide a quantitative measure of the change in the Z(t) sequence brought about by urbanization. Similar analyses are performed by using the other models ML2, MN1 etc. The results from these analyses which are summarized in Table 5.6

are discussed below.

5.3 Results - Functional Series Models

Models ML1 and MN1:

The ratio R_1 in linear and nonlinear models is -0.016 and -0.032 during regeneration, and 0.018 and 0.061 during prediction. Thus the value of R_1 is small in both regeneration and prediction.

The ratio R_2, in linear and nonlinear models is 0.184 and 0.170 during regeneration, and 0.169 and 0.383 during prediction. Thus the linear model gives a lower value of R_2 than the nonlinear model. However the value of R_2 during regeneration is low in both linear and nonlinear models.

The correlation coefficient R_3 is high in linear and nonlinear models during both regeneration and prediction. R_3 values of 0.890 and 0.904 are obtained in regeneration by using linear and nonlinear models. The corresponding values of R_3 in prediction are 0.904 and 0.822. Thus the linear model results in a higher value of R_3 than a nonlinear model during prediction, indicating that the prediction performance of the linear model is better than that of the nonlinear model.

The improvement in the regeneration and prediction performance brought about by incorporating the model for the $e_r(t)$ series into the functional series model may be seen by comparing the results presented in Tables 5.6 and 5.3. The ratio R_2 is reduced from 0.556 to 0.184 when the model ML1 is used whereas it decreases from 0.454 to 0.17 for the models N1 and MN1. Similarly, the correlation coefficient R_3 between the observed and regenerated runoff increases from 0.589 to 0.894 for models L1 and ML1 and from 0.685 to 0.904 for the models N1 and MN1. The value of R_3 computed by using the observed and predicted runoff increases from 0.342 to 0.904 for the models L1 and ML1 and from 0.205 to 0.822 for the nonlinear models N1 and MN1. Similarly the R_1 value is reduced by more than about 8% for the model ML1 and by about 20% for the model MN1. Thus the error model reduces the R_2 value and increases the R_3 value both in linear and nonlinear models.

Models ML2 and MN2:

The ratio R_1, in linear and nonlinear models, is 0.052 and 0.028 during regeneration, and 0.040 and 0.028 during prediction. Thus both linear and nonlinear models give low values of R_1 during regeneration and prediction.

The ratio R_2, is respectively 0.187 and 0.197 in linear and nonlinear models during regeneration, and 0.250 and 0.596 during prediction. Thus linear model gives a lower value of R_2 than nonlinear models during prediction.

The value of R_3 is once again higher in linear model than in nonlinear model. The value of R_3 during regeneration is however high in both linear and nonlinear models.

The error model, when incorporated into the functional series model improves the regeneration and prediction results as seen from 0.722 to 0.187 in the linear model, and from 0.663 to 0.197 in the nonlinear model. During prediction the value of R_2 is reduced from 66.2% to 25% in the linear model and from 131.5% to 59.6% in the nonlinear model. Similarly the value of R_3 in regeneration is increased from 0.370 to 0.877 in the linear model, ML2 and from 0.425 to 0.892 in the nonlinear model MN2. During prediction, the value of R_3 is increased from 0.507 to 0.877 in the linear models, and from 0.001 to 0.767 in the nonlinear model. The error model reduces the value of R_1 by more than about 10% in the linear model and about 16% in the nonlinear model, during both regeneration and prediction.

The regenerated, predicted, and the corresponding observed runoff from the overall models are shown in Fig. 5.3. There is a good correspondence between the regenerated and observed runoff, and between predicted and observed runoff in both linear and nonlinear models. However the prediction errors are smaller in linear models that in nonlinear models. But the peak runoff and times-to-peak runoff are very close to the observed values both in regeneration and prediction for both linear and nonlinear models.

The ratio R_2 remains approximately the same when the models are used to regenerate runoff (0.170 - 0.197). However, except for model ML1, all the other models give poor prediction performance. For example, the ratio R_2 increases from 0.187 in regeneration to 0.25 in prediction. Only the model ML1 gives a *smaller* R_2 value in prediction than in regeneration. This result, however, is suspect as the R_2 value in prediction cannot be less than that in regeneration. In view of the foregoing, we may conclude that the characteristics of the process Z(t) has substantially changed between 1953 and 1963. In fact even the structure of the kernels $H_1(\tau)$ and $H_2(\tau_1,\tau_2)$ also indicate the same result.

5.4 Analysis of the Effects of Urbanization on Runoff Prediction-Nonlinear Stochastic Model

The nonlinear stochastic model discussed in Chapter IV will be used to model the rainfall-runoff process of the Salt Creek basin during the first and second periods. The corresponding valid models will be designated MS1 and MS2. These valid models will be used to predict runoff during the other period by using the precipitation values of that period. In these models the values of R and m have to be specified as discussed in Chapter IV.

5.4.1 Parameter Estimation in the Nonlinear Stochastic Model:

The values of R and m which result in minimum mean square value of error were selected as optimum values of R and m. The ratio R_2 of the error mean square to mean square runoff are plotted in Fig. 5.4 for different values of m and R for the two periods. The results presented in Fig. 5.4 indicate that values of R and m which give minimum mean square error are 10 and 3 for the first period and 15 and 2 for the second period. The total number of parameters are accordingly 44 and 48 respectively

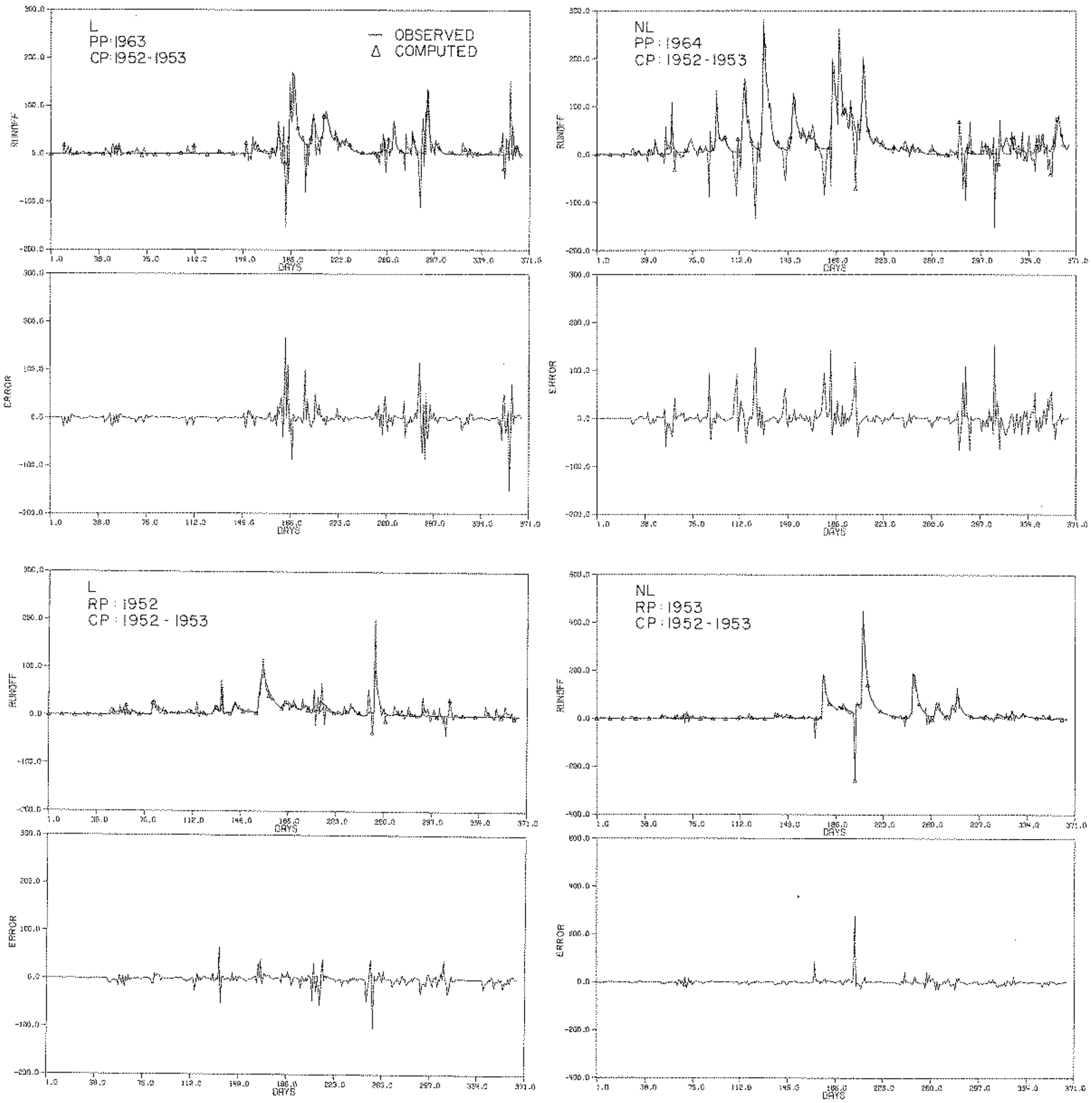


FIG. 5.3 THE OBSERVED, REGENERATED AND PREDICTED RUNOFF FROM THE LINEAR AND NONLINEAR FUNCTIONAL SERIES MODELS

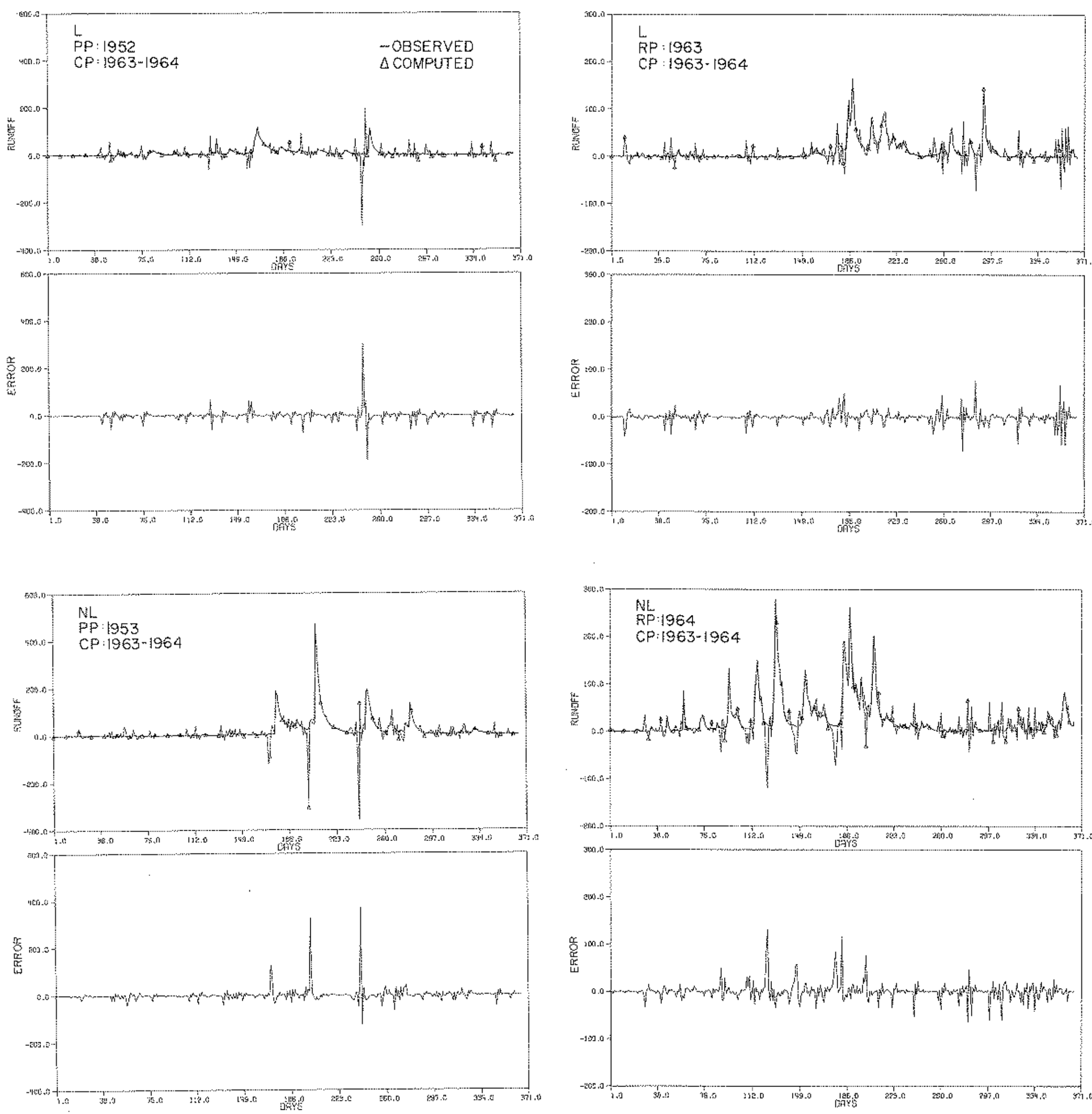


FIG. 5.3 THE OBSERVED, REGENERATED AND PREDICTED RUNOFF FROM THE LINEAR AND NONLINEAR FUNCTIONAL SERIES MODELS (Cont'd)

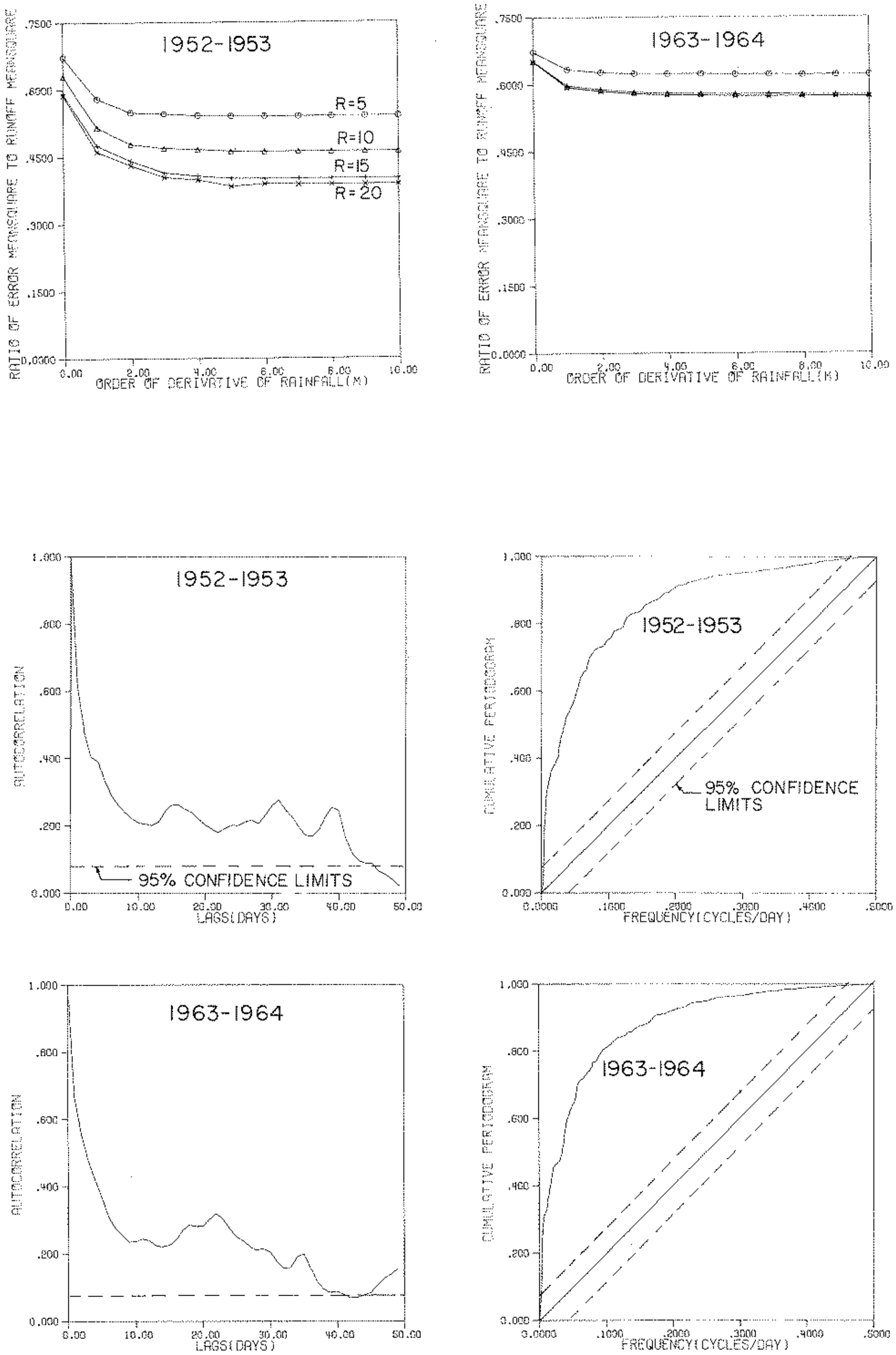


FIG. 5.4 VARIATION OF R_2 WITH R AND m, CORRELOGRAMS, CUMULATIVE PERIODOGRAMS OF REGENERATION ERRORS FROM THE NONLINEAR STOCHASTIC MODELS

for the two periods. The value of m which indicates the highest derivative of rainfall used in modeling the rainfall-runoff process is thus smaller for the second period data than for the first period data, which indicates that there is a faster response of the basin in transforming rainfall into runoff during the second period. The nonlinear models for the two periods can now be written as in eq. 5.11.

First period:

Nonlinear model (S1): $$Q(t) = \sum_{i=0}^{3} \sum_{n=0}^{10} C_n^{(i)} \theta_n^{(i)} \left(\frac{d^i I(t)}{dt^i} \right)$$

Second period: (5.11)

Nonlinear model (S2): $$Q(t) = \sum_{i=0}^{2} \sum_{n=0}^{15} C_n^{(i)} \theta_n^{(i)} \left(\frac{d^i I(t)}{dt^i} \right)$$

The observed runoff Z(t) is related to Q(t) as given in eq. 5.12

$$Z(t) = Q(t) + e_r(t) \tag{5.12}$$

The parameters $C_n^{(i)}$ in the above models were estimated by using the method discussed in Chapter IV, and are given in Table 5.7. These parameters were used to regenerate runoff in the respective periods as discussed in Section 5.2 and the regeneration errors $e_r(t)$ which were computed by using eq. 5.12 were tested for whiteness. The correlograms and the cumulative periodograms of the error sequence $e_r(t)$ are shown in Fig. 5.4 and these indicate that the residuals are correlated and periodic. The results of Portmanteau test given in Table 5.8 also indicate the same result. Although the performance of the models S1 and S2 during regeneration or prediction are expected to be unsatisfactory, the results of regeneration and prediction performance of these are summarized in Table 5.9. These results will be used to obtain an idea of the improvement brought about by the error model. The regeneration and prediction performance of these models are not satisfactory as measured by the values or R_1, R_2 and R_3 given in Table 5.9.

AR models of increasing order were fitted to the errors $e_r(t)$ from models S1 and S2 and it was found that a fifth order AR model was adequate to characterize the error sequence $e_r(t)$. The parameter of AR models were estimated by using the recursive algorithm AL discussed in Chapter IV. The parameter estimates and their standard errors are given in Table 5.10. The parameters β_1, β_2 and β_5 are significant in the error model for the first period while β_1, β_2, β_3 and β_5 are significant for the second period. The error models were incorporated into the nonlinear models S1 and S2 and the performance of the overall model was evaluated during regeneration and prediction. The overall models MS1 and MS2 can now be written as in eqs. 5.13 and 5.14.

TABLE 5.7a PARAMETER ESTIMATES $C_n^{(i)}$ IN THE NONLINEAR STOCHASTIC MODEL (1952-53 DATA)

R↓ m→	0	1	2	3
0	13.640	0.000	0.000	0.000
1	11.347	7.131	2.807	1.109
2	1.545	-0.477	-2.0237	-1.060
3	-8.549	-6.670	-0.701	-0.048
4	-2.581	-1.919	3.553	1.515
5	5.116	4.647	0.300	-0.178
6	2.699	-1.151	-3.718	-1.673
7	-2.813	-4.152	-1.559	1.558
8	-6.091	-1.644	1.127	0.463
9	-1.492	2.379	-0.899	-0.479
10	0.064	1.277	-0.311	-0.229
11	1.436	1.075	-0.006	-1.335
12	2.306	0.670	0.217	-2.303
13	2.961	0.361	0.388	-2.560
14	3.432	0.154	0.370	-2.201
15	3.190	0.003	0.353	-1.609

TABLE 5.7b PARAMETER ESTIMATES $C_n^{(i)}$ IN THE NONLINEAR STOCHASTIC MODEL (1963-64 DATA)

R↓ m→	0	1	2
0	17.701	0.000	0.000
1	10.982	4.958	0.257
2	-3.716	-5.056	-2.358
3	3.514	1.894	1.843
4	2.273	1.837	-0.532
5	0.868	-0.418	-1.092
6	-0.659	-1.649	0.042
7	2.147	1.515	0.2301
8	3.469	0.964	-1.578
9	-3.308	-2.806	-0.013
10	-1.155	1.467	0.468

TABLE 5.8 RESULTS OF *PORTMANTEAU* TEST ON RESIDUALS FROM MODELS S1 AND S2

DATA WATER YEARS	LAG	TEST STATISTIC	CRITICAL VALUE AT $\alpha = 0.05$	DECISION
1952-53	5	675.82	11.07	R
1963-64	5	846.33	11.07	R

TABLE 5.9 REGENERATION AND PREDICTION RESULTS OF MODELS S1 AND S2

DATA (WATER YEARS) AND MODEL	TYPE OF ANALYSIS	ORDER OF RAINFALL DERIVATIVE (m)	MEAN RUNOFF (cfs)	ERROR MEAN SQUARE $\times 10^3$(cfs)^2	ERROR MEAN (cfs)	ERROR MEAN SQUARE $\times 10^3$(cfs)^2	R_1	R_2	R_3
1952-53 Model S1	R	0	13.64	1.294	0.000	0.766	0.000	0.592	0.556
		1	13.64	1.294	0.000	0.615	0.000	0.475	0.667
		2	13.64	1.294	0.000	0.570	0.000	0.515	0.697
		3	13.64	1.294	0.000	0.537	0.000	0.415	0.718
	P	0	17.70	1.423	0.000	1.454	0.000	1.021	0.106
		1	17.70	1.423	0.000	1.454	0.000	1.021	0.106
		2	17.70	1.423	0.000	1.377	0.000	0.967	0.159
		3	17.70	1.423	0.000	1.324	0.000	0.930	0.197
1963-64 Model S2	R	0	17.70	1.423	0.000	0.928	0.000	0.652	0.405
		1	17.70	1.423	0.000	0.852	0.000	0.598	0.483
		2	17.70	1.423	0.000	0.839	0.000	0.589	0.500
	P	0	13.64	1.294	0.000	1.115	0.000	0.895	0.143
		1	13.64	1.294	0.000	1.228	0.000	0.949	0.127
		2	13.64	1.294	0.000	1.289	0.000	0.996	0.097

TABLE 5.10 ESTIMATES OF PARAMETERS IN ERROR MODELS

DATA (WATER YEARS)	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$
1952-53	0.871 (0.056)*	-0.115 (0.074)	$\underline{-0.004}$ $\overline{(0.075)}$	-0.054 (0.074)	0.139 (0.056)
1963-64	0.958 (0.044)	-0.232 (0.061)	0.106 (0.061)	$\underline{-0.024}$ $\overline{(0.061)}$	0.035 (0.004)

First Period:

Model MS1:
$$\hat{Q}(t) = \sum_{i=0}^{3} \sum_{n=0}^{10} C_n^{(i)} \theta_n^{(i)} \left(\frac{d^i I(t)}{dt^i}\right) + \sum_{i=1}^{5} \beta_i \, e_t(t-i) \tag{5.13}$$

Second Period:

Model MS2:
$$\hat{Q}(t) + \sum_{i=0}^{2} \sum_{n=0}^{15} C_n^{(i)} \theta_n^{(i)} \left(\frac{d^i I(t)}{dt^i}\right) + \sum_{i=1}^{5} \beta_i \, e_r(t-i) \tag{5.14}$$

The overall model MS1 for the period 1952-53, and MS2 for the period 1963-64 were used to regenerate and predict runoff in different periods by following the previously discussed procedure. Thus the model MS1 was used to regenerate runoff during 1952-53 and predict runoff during 1963-64. Similarly the model MS2 was used to regenerate runoff during 1963-64 and predict runoff during 1952-53. The regeneration errors $e_r(t)$ from each of these overall models should, however, be uncorrelated, for these models to

*Values in parentheses correspond to the standard error of estimates (Table 5.10).

be valid during the respective periods.

The correlograms and the cumulative periodograms of $e_r(t)$ are shown in Fig. 5.5. The results of *Portmanteau* test on $e_r(t)$ are summarized in Table 5.11. These results indicate that the residual series

TABLE 5.11 RESULTS OF *PORTMANTEAU* TEST ON THE RESIDUALS FROM MODELS MS1 and MS2

DATA (WATER YEARS)	LAG (DAYS)	MODEL	TEST STATISTIC	CRITICAL VALUE AT = 0.05	DECISION
1952-53	40	MS1	51.89	55.76	A
1963-64	50	MS2	61.97	67.50	A

$e_r(t)$ from Models MS1 and MS2 can be considered as uncorrelated upto 50 lags. The histograms of the residuals are shown in Fig. 5.5 indicate that the residuals can be considered approximately normally distributed with zero mean.

The regeneration and prediction results of the overall models for the two periods are given in Table 5.12. The residuals series $e_r(t)$ may be considered to be zero mean in both the models. The

TABLE 5.12 REGENERATION AND PREDICTION RESULTS OF THE MODELS MS1 and MS2

DATA (WATER YEARS)	TYPE OF ANALYSIS	MODEL	MEAN RUNOFF (cfs)	MEAN SQUARE RUNOFF $\times 10^3$(cfs)^2	ERROR MEAN (cfs)	ERROR MEAN SQUARE $\times 10^3$(cfs)^2	R_1	R_2	R_3
1952-53	R	MS1	13.64	1.204	0.027	0.191	0.0	0.172	0.904
	P		17.70	1.423	0.0	0.429	0.0	0.301	0.795
1963-64	R	MS2	17.70	1.423	0.012	0.242	0.0	0.170	0.877
	P		13.64	1.204	0.009	0.523	0.0	0.454	0.726

ratio R_2, is 0.172 and 0.170 for the models MS1 and MS2 during regeneration. The corresponding value of R_2 are 0.301 and 0.404 during prediction. Thus, although the values of R_2 are low during regeneration, the mean square error is larger during prediction. The value of R_3 from these models is lower in prediction (0.795 and 0.740) than during regeneration (0.904 and 0.890) indicating that the prediction errors are larger than the residuals.

The models MS1 and MS2 give better regeneration and prediction performance in comparison with models S1 and S2. The R_2 value in regeneration is reduced by about 25% and 47% in models MS1 and MS2 compared with models S1 and S2. During prediction, R_2 is reduced by more than 60% by models MS1 and MS2. Similarly the R_3 values are increased by 18% and 30% by models MS1 and MS2 in comparison with the values given by models S1 and S2, and the corresponding increases during prediction are about 8% and 7.5% respectively.

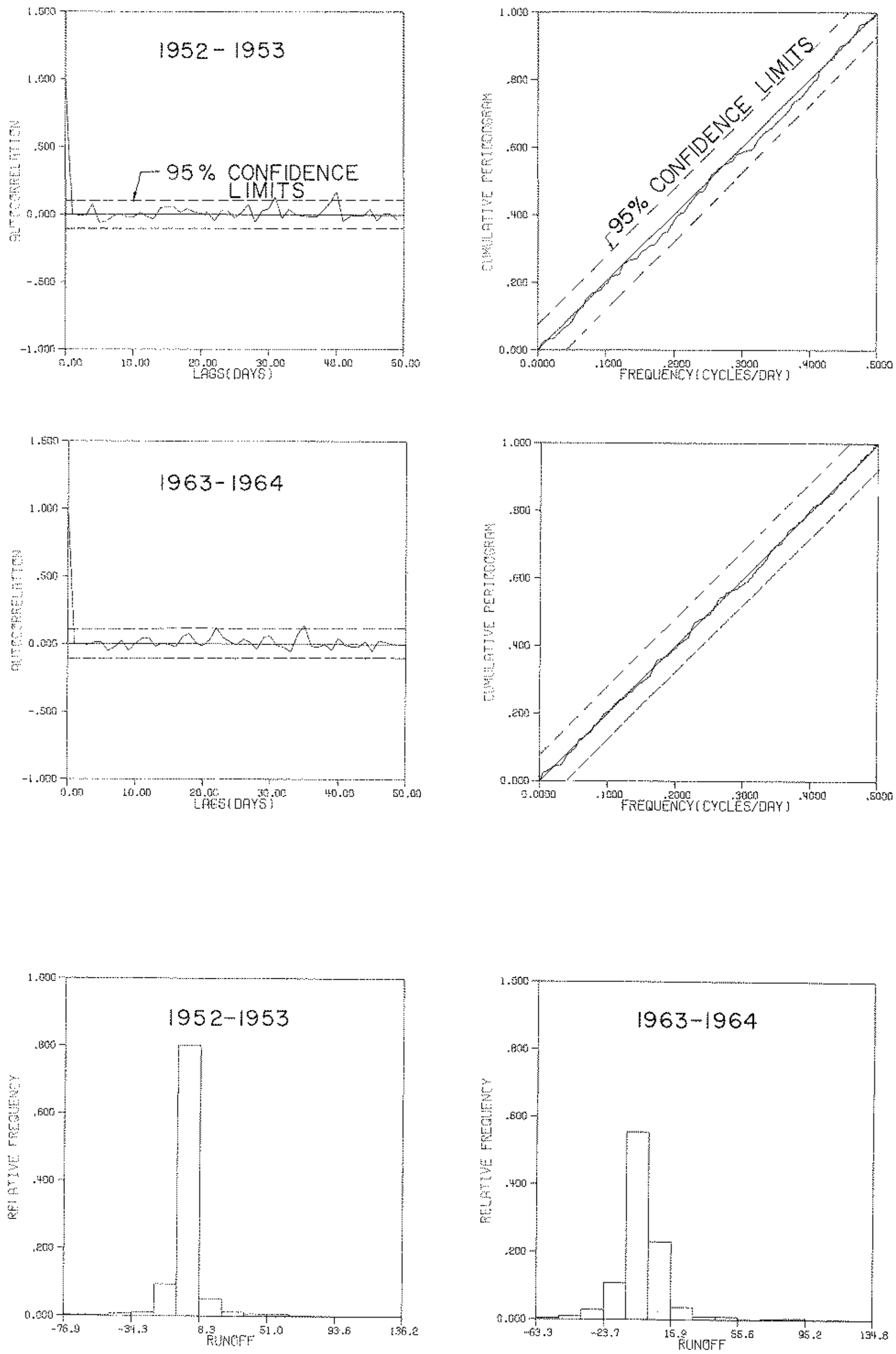


FIG. 5.5 CORRELOGRAMS, CUMULATIVE PERIODOGRAMS AND HISTOGRAMS OF REGENERATION ERRORS FROM THE NONLINEAR STOCHASTIC MODELS

The regenerated and predicted runoff from the overall models are shown in Fig. 5.6. The corresponding observed runoff are also shown in Fig. 5.6. There is a good correspondence between the regenerated and observed runoff in both the periods but considerable deviations are apparent between the predicted and observed runoff. But the peak runoff values and times to peak runoff computed by the models are very close to the corresponding observed values in both regeneration and prediction.

Thus the models MS1 and MS2 can be considered to be valid models during the first and the second period respectively. But when these models are used for predicting runoff during the other period by using the rainfall values of the other period, the results are not satisfactory. The poor prediction performance of these models can be attributed to the change in runoff characteristics due to urbanization. The structure of these valid models and the estimates of parameters are also different for the two periods.

5.5 Simulation of Urban Runoff and the Effects of Urbanization on Runoff

In Section 5.2, valid models were obtained for the rainfall-runoff process for the first and second period data. The regeneration performance of these models were found to be satisfactory during the respective periods, but when used to predict runoff in the other periods by using the corresponding rainfall values, the prediction results were not satisfactory. In this section, the model which is valid for one of the periods is used to generate synthetic runoff in the other period by using the observed rainfall values of that period and vice versa. The simulation capabilities of these models will be first established and then the simulation results will be used to analyze the effect or urbanization on runoff.

5.5.1 Generation of Synthetic Runoff:

The overall linear and nonlinear models given in eq. 5.8 and the nonlinear stochastic models in eqs. 5.13 and 5.14 were used to generate synthetic runoff during the two periods. The models ML1, MN1 and MS1 were used to generate the runoff sequence in the second period and the models ML2, MN2, MS2 were used to generate data in the first period. The runoff values were generated from these models by using eq. 5.15.

$$\hat{Q}_G(t) = \hat{Q}(t) + W(t) \tag{5.15}$$

where $\hat{Q}(t)$ is the runoff from the overall models given in eqs. 5.8, 5.13 and 5.14 and $W(t)$ is defined in eq. 5.7. The $W(t)$ values in these models were assumed to be (corresponding to the residuals indicated by R) normally distributed with the zero mean and variance given in Table 5.6 and 5.12. The difference between the generated and predicted runoff is that the noise term is included in generating runoff while it is not considered in predicting runoff. The first fiew values of $e_r(t-1)$ were assumed to be zero. The generated and the observed runoff in the two periods are plotted in Fig. 5.7. It can

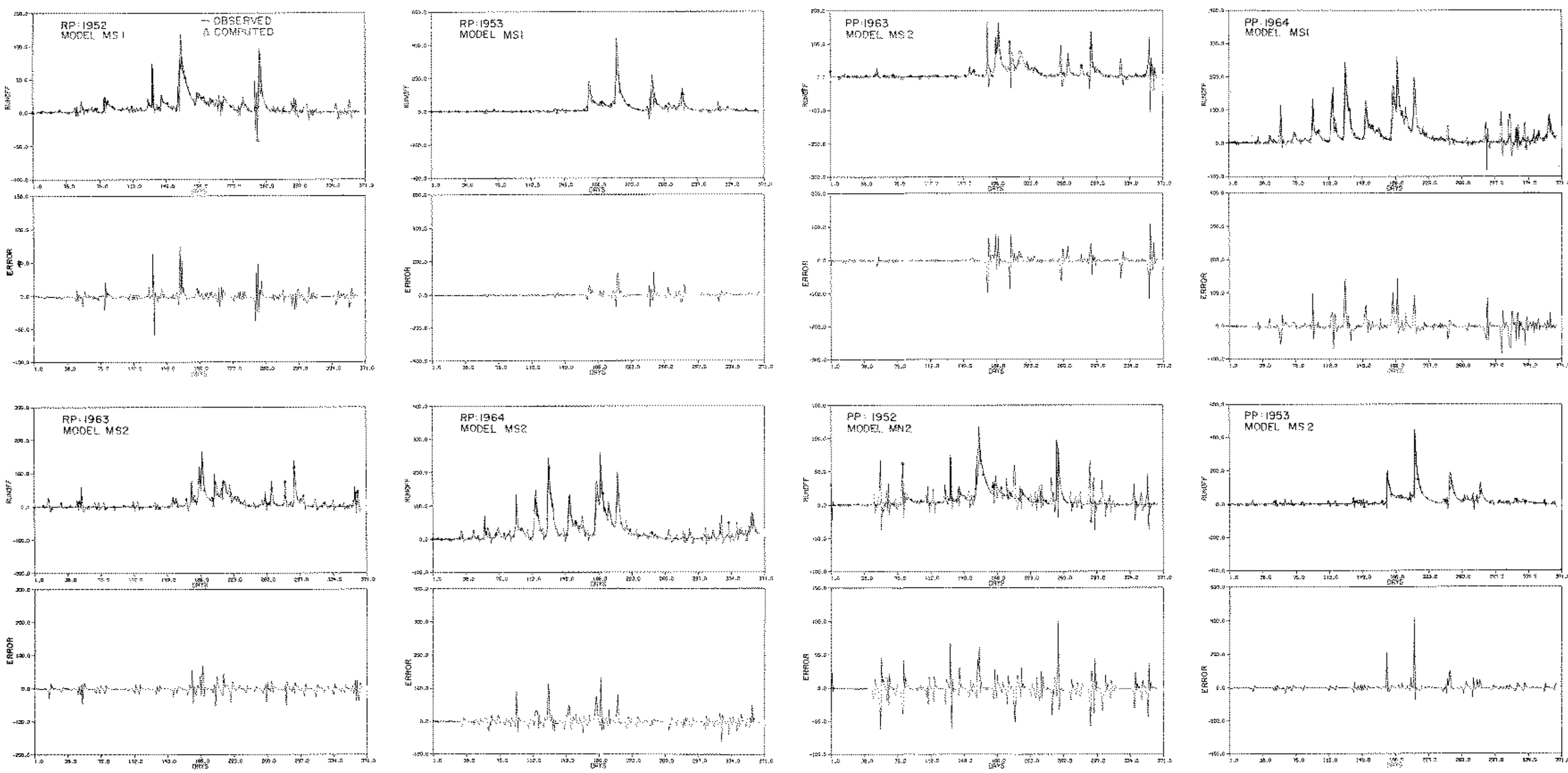


FIG. 5.6 OBSERVED, REGENERATED AND PREDICTED RUNOFF FROM THE NONLINEAR STOCHASTIC MODELS

be seen from Fig. 5.7 that the peak values and time-to-peak values of observed and generated runoff area very close to each other although the correspondence between smaller values of observed and generated runoff is not very good. Several properties of the generated sequences were comapred with those of the observed sequences and these properties are given below.

1. Basic statistics of runoff
2. Histograms of runoff
3. Correlograms of runoff
4. Mass curves of runoff
5. Double mass curves of rainfall and runoff
6. Flow duration curves
7. Frequency analysis of extreme values of runoff
8. Storm analysis

The above characteristics were computed for the generated and observed runoff in the two periods and are discussed below.

5.5.2 Simulation Results

1. *Basic statistics of runoff:*

The basic statistics of the generated and the observed runoff in the two periods are given in Table 5.13. The mean, variance, skewness coefficient and the maximum value of the generated and observed runoff are very close to each other in all the models.

2. *Histograms of runoff:*

The histograms of the generated and observed runoff in the two periods are shown in Fig. 5.8 for all the models. The histograms of generated sequences and those of the observed sequences are very clsoe to each other in all the models.

3. *Correlograms of runoff:*

The correlograms of the generated and the observed runoff are shown in Fig. 5.8. The correlograms of the generated runoff from all the models roughly agree with the corresponding correlograms of observed runoff. The correlograms of the generated data of the first period, is closer to the correlograms of the observed data than the correlograms of the second period.

4. *Flow mass curves:*

The mass curves of the generated and observed data are shown in Fig. 5.9. The observed and generated runoff mass curves are very close to each other in both the periods and for all the models. Also the mass curve of the generated runoff of 1963-64 period plots to the left of the mass curve of the generated runoff during 1952-53, thereby indicating that the runoff has increased during 1963-64.

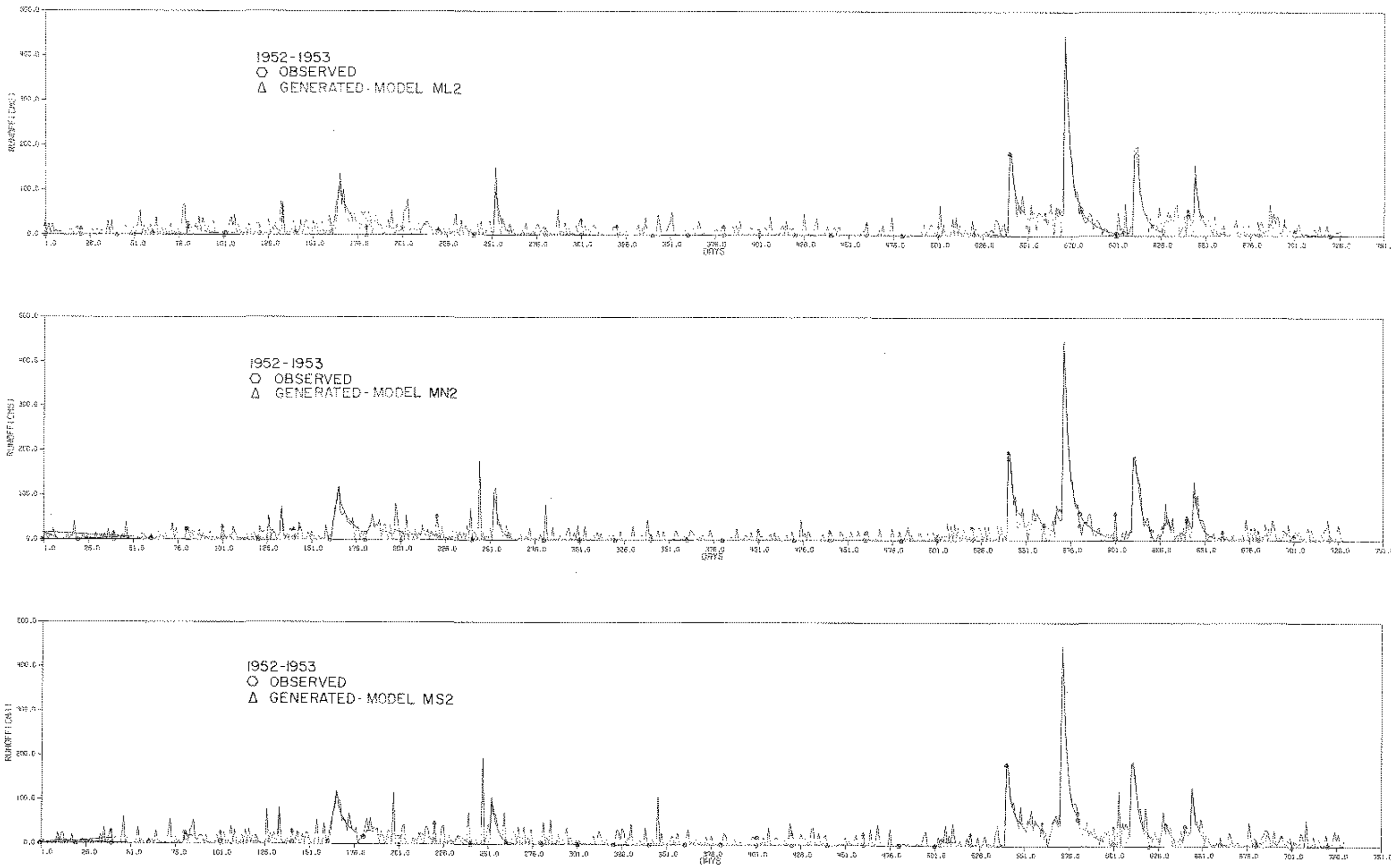


FIG. 5.7 OBSERVED AND GENERATED RUNOFF FROM THE FUNCTIONAL SERIES AND NONLINEAR STOCHASTIC MODELS

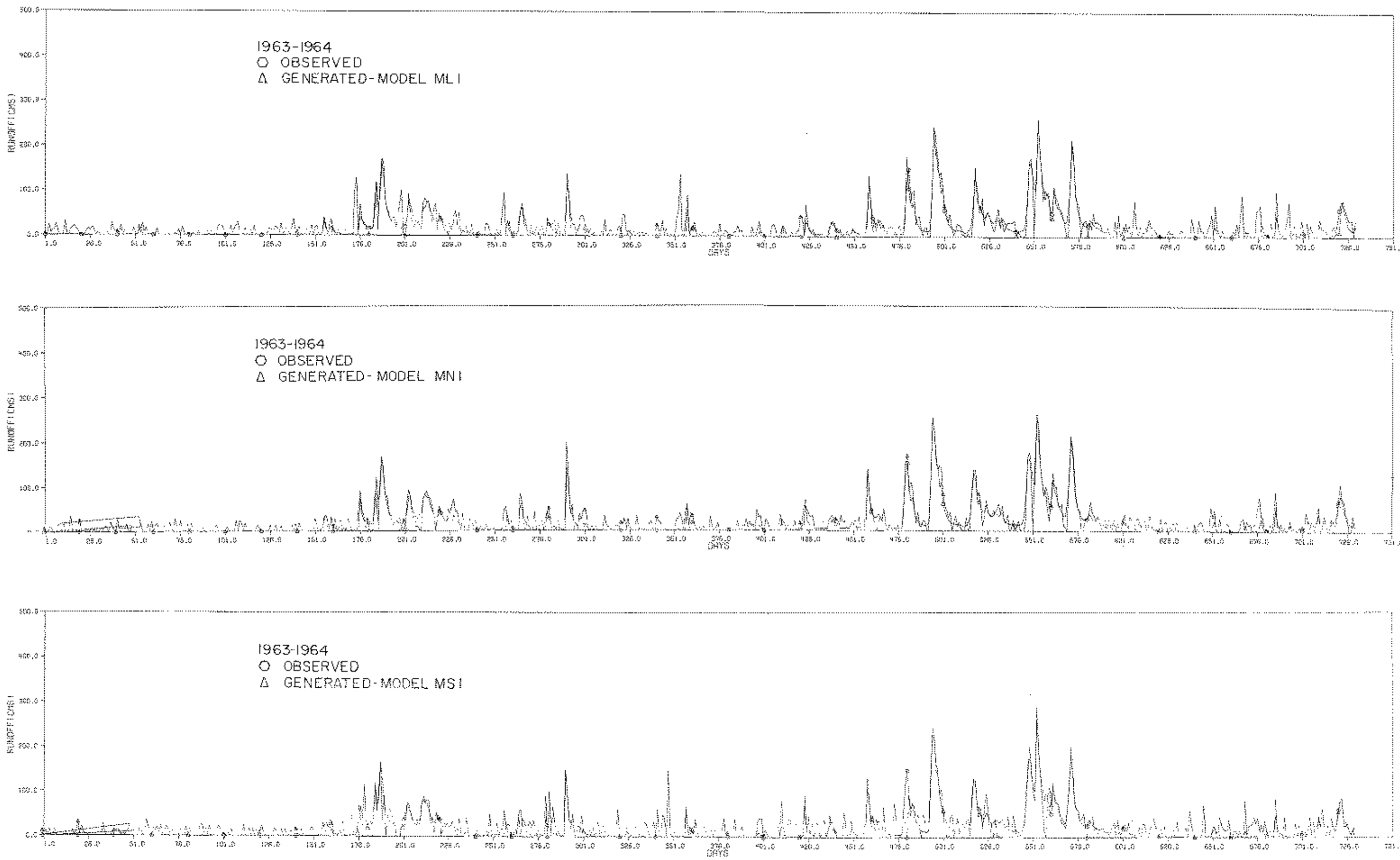


FIG. 5.7 OBSERVED AND GENERATED RUNOFF FROM THE FUNCTIONAL SERIES AND NONLINEAR STOCHASTIC MODELS (Cont'd)

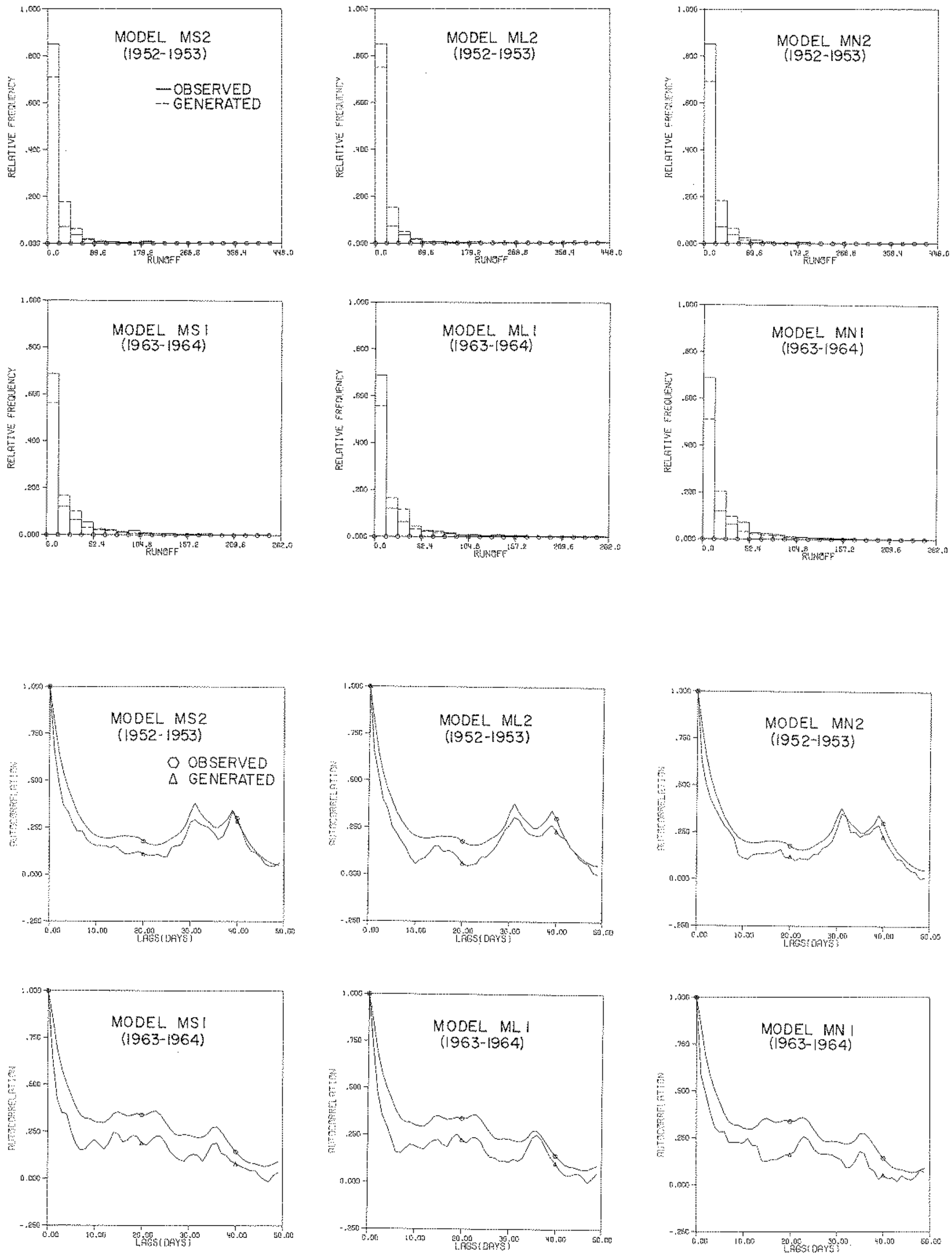


FIG. 5.8 HISTOGRAMS AND CORRELOGRAMS OF THE OBSERVED AND GENERATED RUNOFF

TABLE 5.13 BASIC STATISTICS OF GENERATED AND OBSERVED DATA

DURATION (WATER YEARS)	MODEL	MEAN	VARIANCE	SKEWNESS COEFF.	MAX. VALUE
1952-53	obs	13.641	1109.2	6.634	448.00
	MS2	13.810	1212.1	5.582	416.80
	ML2	12.320	1212.1	5.592	431.00
	MN2	14.581	1319.0	4.176	428.90
1963-64	obs	17.701	1112.0	3.392	262.00
	MS1	17.480	1221.5	2.892	247.80
	ML1	18.230	1300.1	3.027	253.70
	MN1	17.694	1173.2	2.921	290.80

5. *Double mass curves:*

The double mass curves of the generated runoff and observed rainfall are shown in Fig. 5.9. The double mass curves of the observed runoff and rainfall are also shown in Fig. 5.9. These double mass curves are close to each other. Also the double mass curves during 1963-64 plot to the left of those during 1952-53, indicating increased runoff during 1963-64 as compared to the runoff during 1952-53.

6. *Flow duration curves:*

The flow duration curves of the generated runoff in the two periods are shown in Fig. 5.10. The flow duration curves of the observed runoff in these periods are also shown in Fig. 5.10. The flow duration curves of generated and observed runoff are very close to each other for all models. It can also be seen from Fig. 5.10 that considerable deviation exists in the middle region of the curves during 1952-53 and 1963-64. The pronounced deviation in the low flow characteristics (up to about 100 cfs) in the two periods is also evident from Fig. 5.10. The duration curves indicate perceptible 90% flows, and the 90% flow during 1963-64 is higher than that during 1952-53. Similarly the flow equaled or exceeded 20% of the time is higher during 1963-64 than during 1952-53. The runoff has thus increased during 1963-64 as compared to the runoff during 1952-53.

7. *Frequency analysis of exceedence series:*

In the present study a frequency analysis of the exceedence series of the observed and generated runoff series was conducted. The partial duration series, is a series of data which are so selected that their magnitude is greater than a certain value. The base value was selected such that the resulting exceedence series is independent. The base values selected for the observed and the generated runoff are given in Table 5.14. Gumbel's type I distribution was fitted to the exceedence series.

The flood frequency curves of the exceedence series during the first and second period are shown in Fig. 5.11 for both the observed and generated runoff. The magnitude of the events against the reduced variate (*Gumbel, 1954*) and also against the return period are shown in Fig. 5.11. The fitted

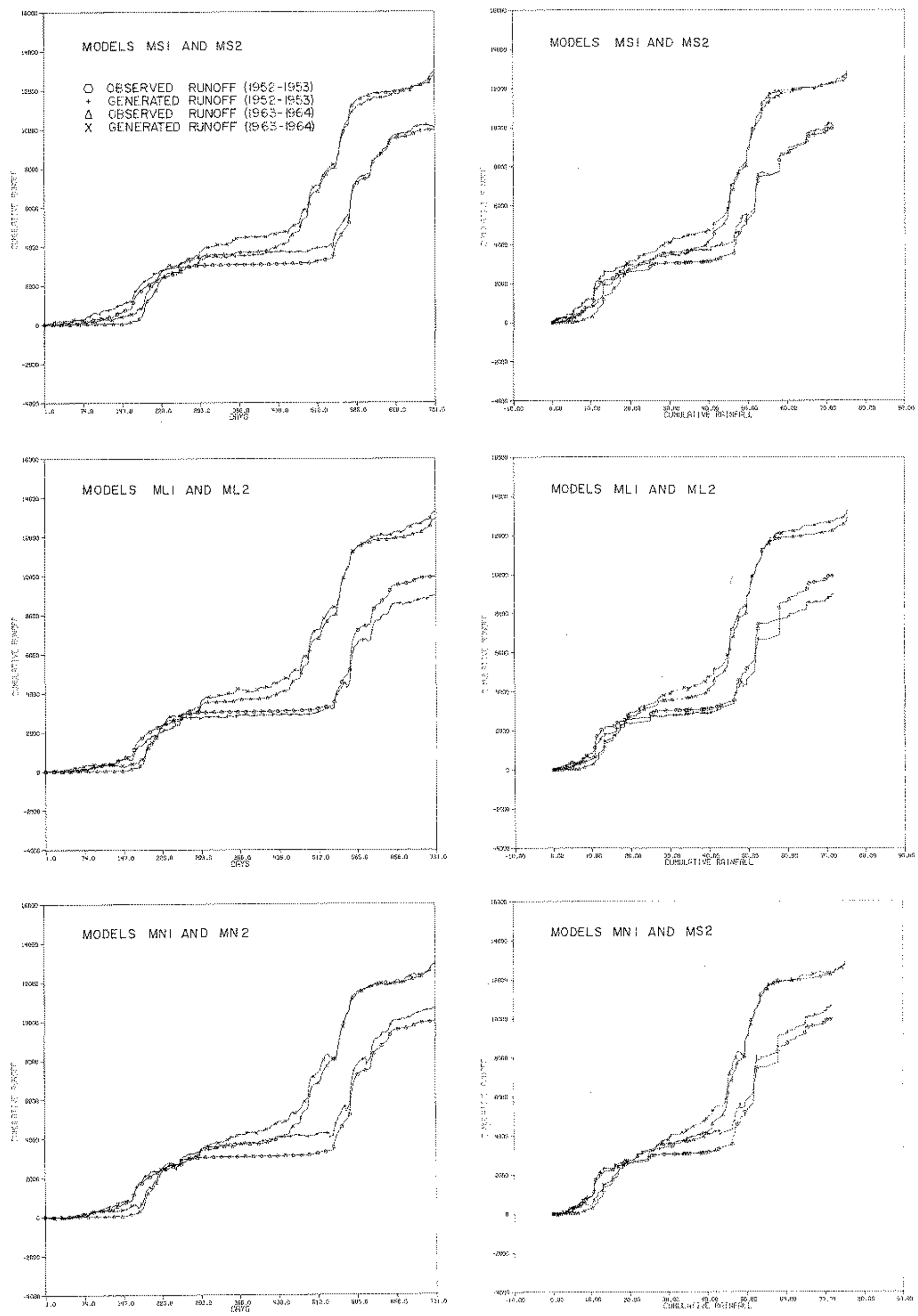


FIG. 5.9 MASS CURVES OF THE OBSERVED AND GENERATED RUNOFF

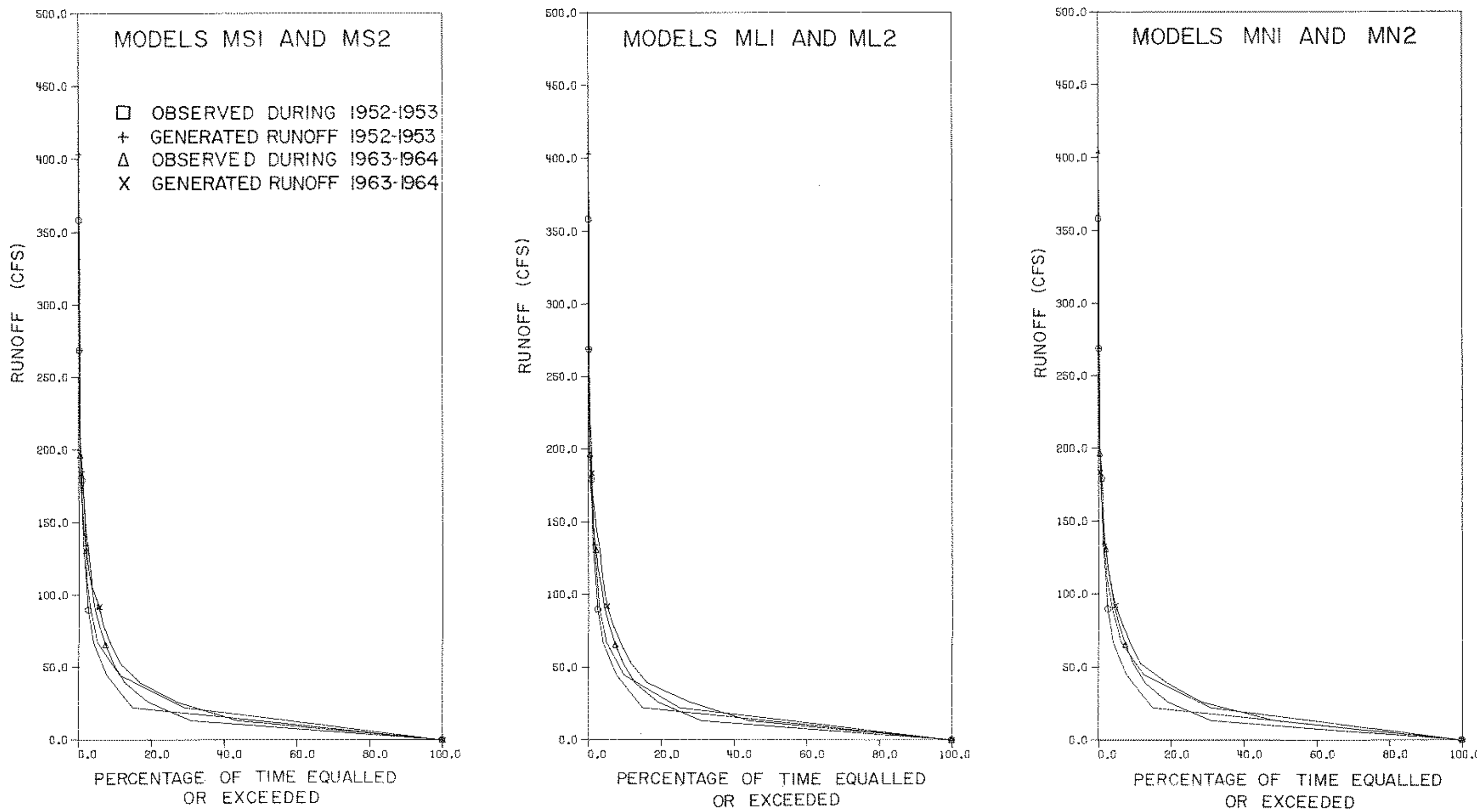


FIG. 5.10 FLOW DURATION CURVES OF THE OBSERVED AND GENERATED RUNOFF

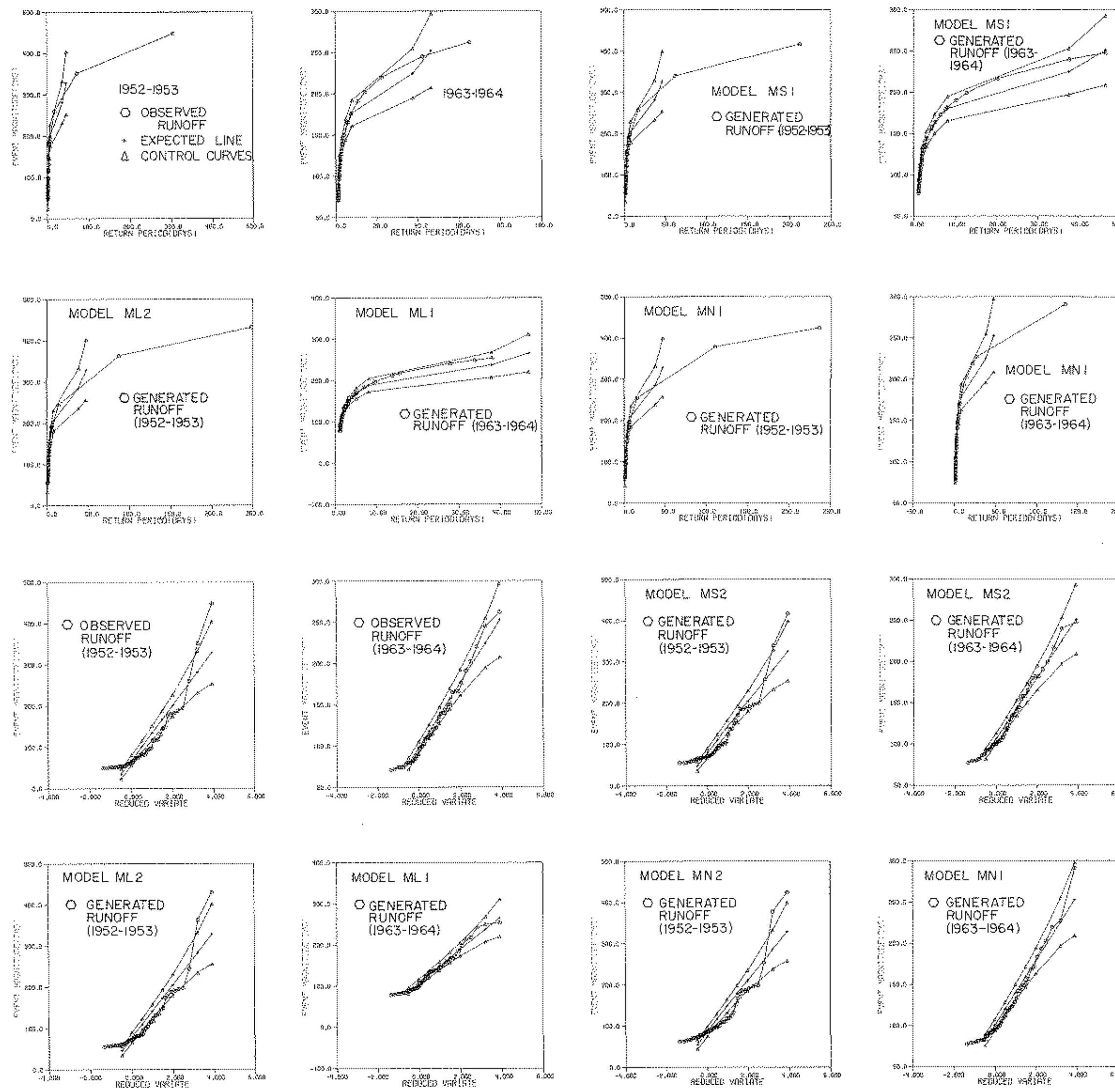


FIG. 5.11 FREQUENCY CURVES OF EXCEEDENCE SERIES OF THE OBSERVED AND GENERATED RUNOFF

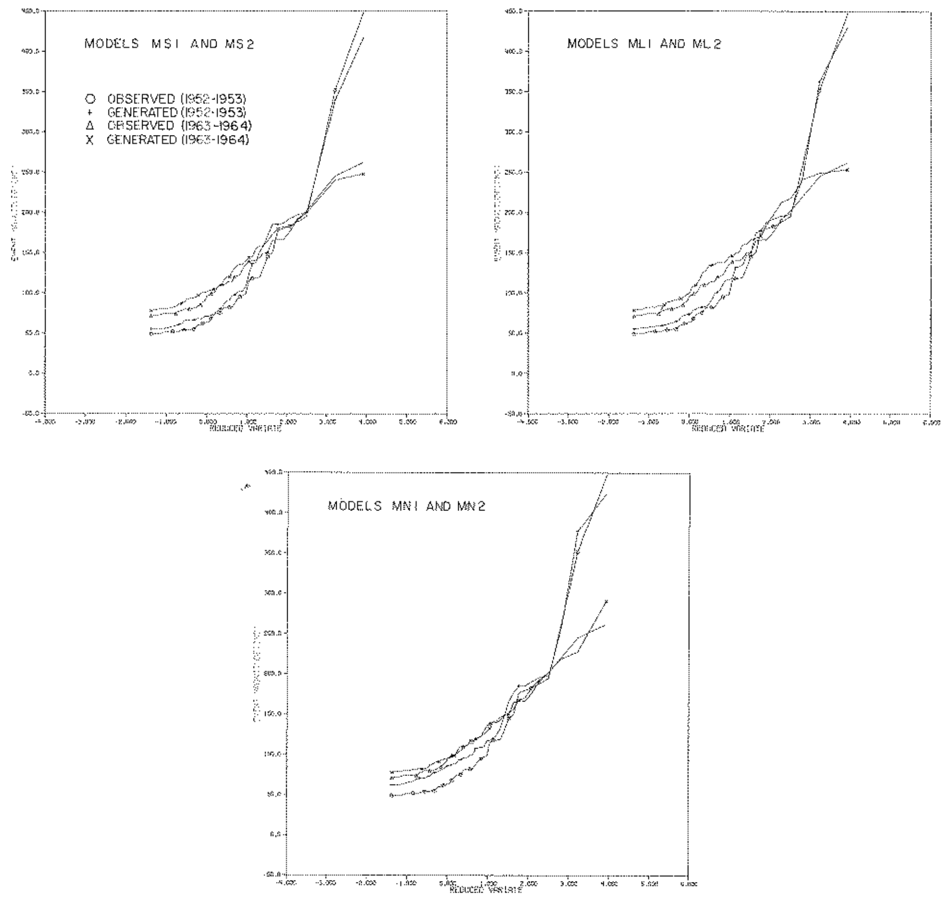


FIG. 5.11 FREQUENCY CURVES OF EXCEEDENCE SERIES OF THE OBSERVED AND GENERATED RUNOFF (Cont'd)

TABLE 5.14 ESTIMATES OF u AND $1/\alpha$ IN GUMBEL'S EQUATION

DURATION OF DATA	MODEL OR DATA	u	$1/\alpha$	Q*
1952-54	Observed	68.24	66.44	68.24
	ML2	77.77	64.09	77.77
	MN2	78.35	63.16	78.35
	MS2	85.43	62.14	85.43
1963-64	Observed	98.03	39.33	98.00
	ML1	108.92	39.95	108.92
	MN1	106.68	36.82	106.68
	MS1	100.53	38.86	100.53

*Q = discharge used to compute exceedence series.

line and the control curves of the theoretical distribution function for the exceedence series are also shown in Fig. 5.11. The frequency distribution of observed and generated runoff exceedence series for the two periods are shown in Fig. 5.11 for each of the models. It can be seen from Fig. 5.11 that the flood frequency curves of the exceedence series of the generated and the observed runoff sequences are very close in the two periods and for all models. The frequency curves for the first period plots to the left of those for the second period, thereby indicating that for a given probability (or the reciprocal of return period), the magnitude of the floods are larger in the second period than in the first period. The generated runoff from all the three models indicate the same result.

8. *Storm analysis:*

Several individual storms were selected by using the observed rainfall and runoff sequences in the two periods and the corresponding rainfall volumes were computed for the observed data. The corresponding runoff volumes of the observed and generated runoff sequences were also computed. The storm runoff volumes of the observed and generated runoff were very close in the two periods and for all models. The flood runoff volumes of the generated data were in good agreement with those of the observed data and for the same rainfall value larger storm runoff volumes are obtained during the second period compared to the first period. Such larger storm runoff volumes are more frequent in the second period than in the first period. Consequently, the urbanization has increased the magnitude and frequency of occurrence of storm runoff.

The peak runoff of each of the selected storms and the corresponding rainfall volume were computed for both the observed and generated data. The peak values of the generated series are approximately close to those of the observed series.

5.6 Conclusions

The various characteristics of runoff discussed in this section indicate that there is a good correspondence between the observed and generated runoff. The simulation capabilities of all the three types of models, linear, second order and the nonlinear stochastic models are approximately the same. The generated runoff in the two periods indicate a change in the runoff characteristics such as the maximum value, skewness coefficient and correlograms. An increase in runoff is indicated by the mass and flow duration curves also. The exceedence analysis of floods shows that the flood flows of higher magnitude are more frequent during the second period than those during the first period. The storm runoff volumes during the second period is generally higher than those during the first period.

CHAPTER VI

SUMMARY AND CONCLUSIONS

Urbanization changes the runoff characteristics of a basin significantly. In the present study the effects of urbanization on runoff has been investigated by developing new methods of modelling the rainfall-runoff process, and by analyzing the prediction and simulation results from these models. Daily rainfall and runoff data (of water years 1952-53 and 1963-64) from the Salt Creek basin, Illinois were used to evaluate the effects of urbanization on runoff from the basin. The effects of urbanization on runoff were analyzed by using some new methods of identification of rainfall-runoff models. The functional series and nonlinear stochastic models of rainfall runoff process were used in the analysis.

The functional series models can be formulated as a linear or higher order nonlinear system model. A recursive algorithm is proposed to estimate the kernel functions of linear or nonlinear functional series models. The algorithm is easy to program and requires small computer storage. It is particularly useful for analyzing long records of data and it can also be used to update the estimates of kernels. The regeneration and prediction performances of the functional series model in modelling the Salt Creek basin rainfall-runoff process is shown to be satisfactory.

In the nonlinear stochastic model, the observed rainfall and its time derivatives are used as inputs. The optimum output of the model involves the probabilistic characteristics of the rainfall and runoff, and the time derivative of rainfall sequences. The coefficients of the system equations are estimated by using the statistical moments of the rainfall and runoff. The accuracy of prediction of runoff can be improved in this model by using higher order derivatives of the rainfall sequence. Thus a new method of modelling rainfall-runoff process has been developed in which the rainfall sequence and its derivatives are used as inputs. The estimation procedure requires small computer storage and does not require large computer time, even for long records of rainfall-runoff data, such as the daily data. The method was once again used to model the rainfall-runoff process of Salt Creek basin and the method gives satisfactory regeneration and prediction results.

After obtaining the coefficients of a model, it is necessary to test the model, in the sense that the regeneration errors have to be uncorrelated and free from periodicities. Otherwise, not only the model is not valid, but the results from the model will be poor in regeneration and prediction. This aspect has been demonstrated in both the above discussed models.

In analyzing the effects of urbanization on runoff characteristics of Salt Creek basin, the following procedure was adopted. Both first and second order functional series and nonlinear stochastic models were used. Valid models were obtained for the two periods. These models were designated ML1, MN1 and MS1 for the first period and ML2, MN2, MS2 for the second period. The valid model for the first period was then used to predict runoff in the second period by using the rainfall values of the second period. Similarly the valid model for the second period was used to predict runoff in the first period. The prediction results indicate that a model valid in one period does not perform well in predicting runoff in the other period. These results thus indicate the changes in runoff characteristics brought about by urbanization of the basin.

The effects of urbanization on runoff characteristics were further investigated by simulating runoff in the two periods by using the valid models. The valid first period model was used to generate synthetic runoff in the second period and vice versa. Several properties of the generated and observed runoff in the two periods were compared. These properties included the flow mass curves, the flow duration curves, the frequency curves of exceedence series, peak analysis, and basic statistical characteristics including histograms, correlograms. The main results from this analysis are presented below.

The characteristics of the generated and observed runoff in the two periods are very close in all the models - linear and second order functional series, and nonlinear stochastic models. Thus the simulation capabilities of the new methods of modelling rainfall-runoff is demonstrated.

The comparison of generated runoff in the two periods indicate that there have been changes in the runoff characteristics of the Salt Creek basin, which can be attributed to the effects of urbanization. The generated runoff in the two periods indicate a change in the runoff characteristics such as the extreme values, skewness coefficient and correlograms. An increase in runoff volume is indicated by the mass and flow duration curves also. The exceedence analysis of floods indicate that the flood flows of higher magnitude are more frequent during the second period than those during the first period. The storm runoff volumes during the second period are generally higher than those during the first period.

Based on these results, the following conclusions can be made

1) A recursive algorithm is proposed for the estimation of kernels in the functional series model of rainfall-runoff process. The estimation procedure is computationally simple and has been shown to give good prediction and simulation results.

2) A nonlinear stochastic model of rainfall-runoff process is developed in which time derivatives of rainfall sequence are used as additional inputs to the model. The estimation procedure has been demonstrated to give good prediction and simulation results.

3) The effects of urbanization on runoff characteristics of an urbanizing basin can be investigated by using the nonlinear models of rainfall-runoff process. Both prediction and simulation results of these models can be used to analyze the effects of urbanization quantitatively.

4) The urbanization of watersheds drastically alters the basin response.

CHAPTER VII

LIST OF REFERENCES

1. Anderson, R. L., "Distribution of the Serial Correlation Coefficient," Ann. Math. Stat., Vol. 13, No. 1, 1942.

2. Anderson, D. G., "Effects of Urban Development on Floods in Northern Virginia," open file Report, U.S. Geol. Sur., Water Resources Div., Richmond, Va., 1968.

3. Bidwell, V. J., "Regression Analysis of Nonlinear Catchment Systems," Water Resources Research, Vol. 7, No. 5, 1971.

4. Boneh, A., and M. H. Diskin, "Demonstration of Nonlinear Effects of a Second Order Surface Runoff Model," Proc. of the Second International Symp. in Hydrology, "Floods and Droughts," pp. 157-166, Water Resources Publications, Ft. Collins, Colorado, 1972.

5. Bose, A. G., "A Theory of Nonlinear Systems," M.I.T. Res. Lab., Electron., Tech. Rep. No. 309, M.I.T., Cambridge, Massachussetts, 1956.

6. Box, G. E. P., and G. M. Jenkins, Time Series Analysis, Forecasting and Control, Holden Day Publ. Co., San Fransisco, California, 1970.

7. Brandstetter, A., and J. Amorocho, "Generalized Analysis of Small Watershed Responses," Water Science and Engineering Paper No. 1035, Department of Water Science and Engineering, University of California, Davis, California, 1970.

8. Bras, R. L. and F. E. Perkins, "Effects of Urbanization on Catchment Response," Journal of the Hydraulics Div., Proc. ASCE, Vol. 101, No. HY3, 1975.

9. Brilliant, M. B., "Theory of the Analysis of Nonlinear Systems," Mass. Inst. Technol., Res. Lab., Electron., Tech. Rep. No. 345, 1958.

10. Carter, R. W., "Magnitude and Frequency of Floods in Suburban Areas," U.S. Geol. Sur., Professional paper 424-B, Art. 5., pp. 9-11, 1961.

11. Chow, V. T., (Ed), Handbook of Applied Hydrology, McGraw Hill, New York, N.Y.

12. Espey, W. H., Jr., C. W. Morgan, and F. D. Masch, "A Study of Some Effects of Urbanization on Storm Runoff From a Small Watershed," Report No. 23, Texas Water Development Board, Austin, Texas, August 1966.

13. Espey, W. H., Jr., C. W. Morgan, and F. D. Masch, "A Study of Some Effects of Urbanization on Storm Runoff From a Small Watershed," Tech. Report No. HYD07-6501, CRWR-2, Center for Research in Water Resources, The Univ. of Texas, Austin, Texas, July 1965.

14. George, D. A., "Continuous Nonlinear Systems," Mass. Inst. Technol. Res. Lab., Electron., Tech., Rep. No. 355, 1959.

15. Gumbel, E. J., "Statistical Theory of Extreme Values and Some Practical Applications," Natl. Bur. Standards (U.S.) Appl. Math. Ser. 33, February 1954.

16. Harder, J. A., and S. M. Zand, "The Identification of Nonlinear Hydrologic Systems," Tech. Rep. No. HEL-8-2, Hydraulic Engineering Lab., University of California, Berkeley, California, 1969.

17. Hino, M., Sukigara, T., and T. Kikkaya, "Nonlinear Runoff Kernels of Hydrologic Systems," Proc. 1st U.S.-Japan Bilateral Seminar in Hydrology, "Systems Approach to Hydrology," Water Resources Publications, Ft. Collins, Colorado, 102-112, 1971.

18. Hubbell, P. G., "Digital Identification and Simulation of Nonlinear Systems with Physiological Applications," Ph.D. thesis, University Microfilms Inc., Ann Arbor, Michigan, 1967.

19. Jacoby, S. L. S., "A Mathematical Model for Nonlinear Hydrologic Systems," J. Geophys. Res., 71(20), 4811-4824, 1966.

20. Kashyap, R. L., and A. R. Rao, Dynamic Stochastic Models From Empirical Data, Academic Press Inc., New York, N.Y. 1976.

21. Kulandaiswamy, V. C., "A Basic Study of the Rainfall Excess-Surface Runoff Relationship in a Basin System," Ph.D. Thesis, Univ. of Illinois, Urbana, Illinois, 1964.

22. Kutchment, L. S., and E. N. Borschchevsky, "Identification of Nonlinear Hydrologic Systems," Meteorol, Gidrol., No. 1, 42-47, 1971.

23. Lee, Y. W., and M. Schetzen, "Measurement of the Wiener Kernels of a Nonlinear System by Cross Correlations," International J. of Control, 2(3), 237-254, 1956.

24. Rao, A. R., and R. G. Srinivasa Rao, "Comparative Analysis of Estimation Methods in Nonlinear Functional Models of the Rainfall-Runoff Process," Tech. Rep. No. 56, Purdue University Water Resources Research Center, West Lafayette, Indiana, 1974.

25. Rao, A. R., J. W. Delleur, and P. B. S. Sarma, "Conceptual Hydrologic Models for Urbanizing Basins, J. of the Hydraulics Div., A.S.D.E., Vol. 98, No. HY7, pp. 1205-1220, 1972.

26. Rao, A. R., and R. L. Kashyap, "Analysis, Construction and Validation of Stochastic Models for Monthly River Flows," Tech. Rept. CE-HYD-73-1, School of Civil Engineering, Purdue University, W. Lafayette, In., 1973.

27. Rao, A. R., R. G. Srinivasa Rao, and R. L. Kashyap, "Stochastic Models for the Ground Water Levels," Tech. Rep. No. 67, Purdue University Water Resources Research Center, West Lafayette, Indiana, 1975.

28. Rao, A. R., and R. G. Srinivasa Rao, "Analysis of the Effects of Urbanization on Rainfall Characteristics I, Tech. Rep. No. 50, Purdue Univeristy Water Resources Research Center, West Lafayette, Indiana, 1974.

29. Riley, J. P., and V. V. Dhruvanarayana, "Application of an Electronic Analog Computer to the Evaluation of the Effect of Urbanization on the Runoff Characteristics of Small Watersheds," Int. Ass. Sci. Hydrology, Pub. 80, Vol. 1, pp. 38-48, December 1968.

30. Salvadori, M. G., and M. L. Baron, Numerical Methods in Engineering, Prentice-Hall Inc., 1961.

31. Savini, J., and J. C. Kammerer, "Urban Growth and the Water Regimen," U.S. Geol. Sur., Water Supply Paper 1591-A, 1961.

32. Speiker, A. M., "Water in Urban Planning, Salt Creek basin, Illinois," U.S.G.S. Water Supply paper No. 2002, U.S. Department of the Interior, Washington, 1970.

33. Srinivasa Rao R. G., "Prediction of Daily Runoff and Evaluation of the Effects of Urbanization on Runoff Characteristics by Using Nonlinear Rainfall-Runoff Models," Ph.D. thesis, Purdue University, West Lafayette, Indiana, December 1975.

34. Stall, J. B., and H. F. Smith, "A Comparison of Urban and Rural Runoff," paper presented at the Hydraulics Div. Conferences of the Am. Soc. Civil Engrs., in Urbana, Ill., August 1961.

35. Tao, P. C., A. R. Rao, and C. Ruckvichai, "Stochastic Forecasting Models of Reservoir Inflows for Daily Operation-I," Tech. Rep. School of Civil Engineering, Purdue University, West Lafayette, Indiana, 1975.

36. Vansickle, D., "The Effects of Urban Development on Storm Runoff," The Texas Engineer, Vol. 32, No. 12, December 1962.

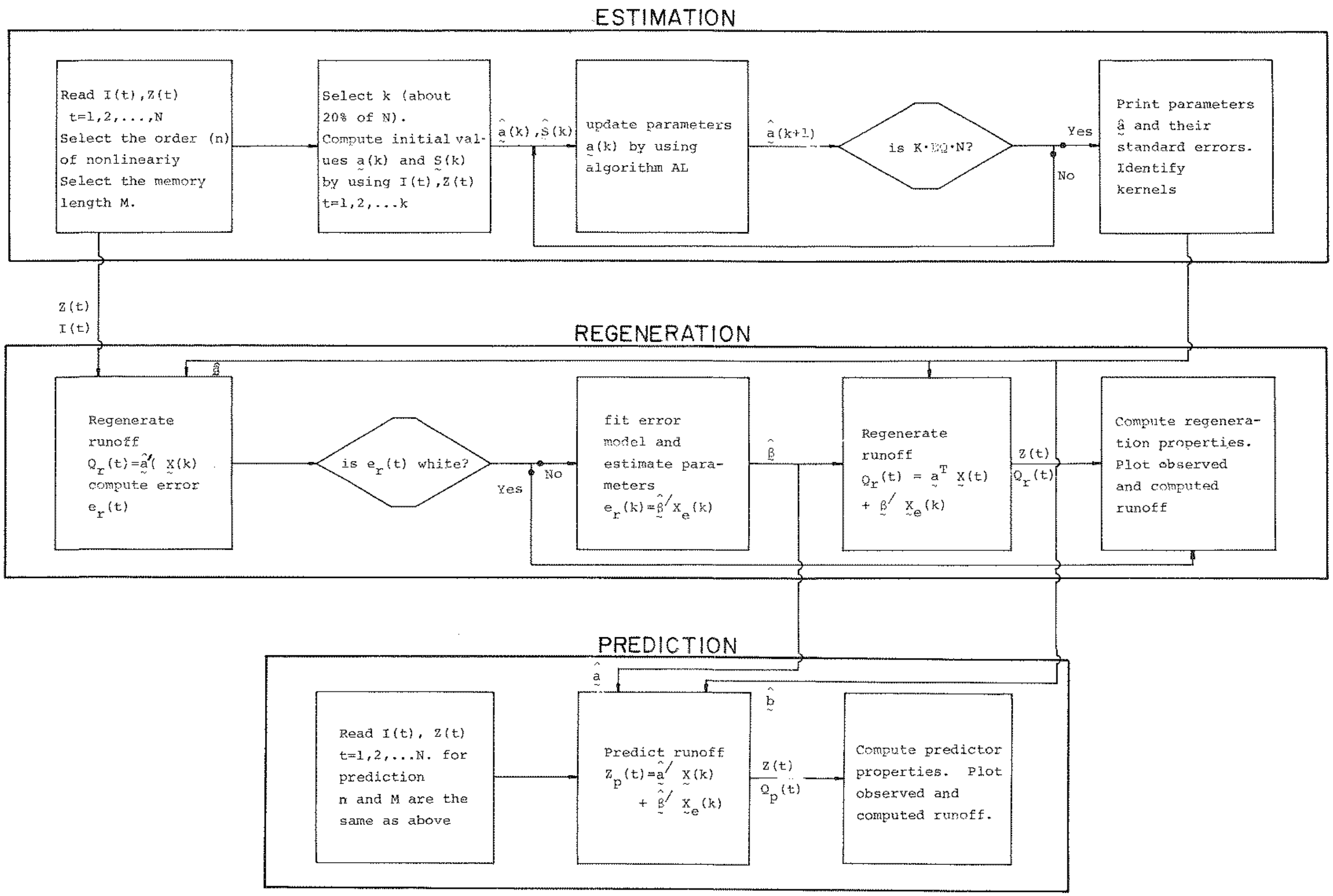
37. Wiener, N., "Response of a Nonlinear Device to Noise," Radiation Lab. Rep. 129, M.I.T., Cambridge, Massachussetts, 1942.

38. Wiitala, S. W., "Some Aspects of the Effect of Urban and Suburban Development Upon Runoff," open file Report, U.S. Geol. Sur., Water Resources Div., Lansing, Mich., 1961.

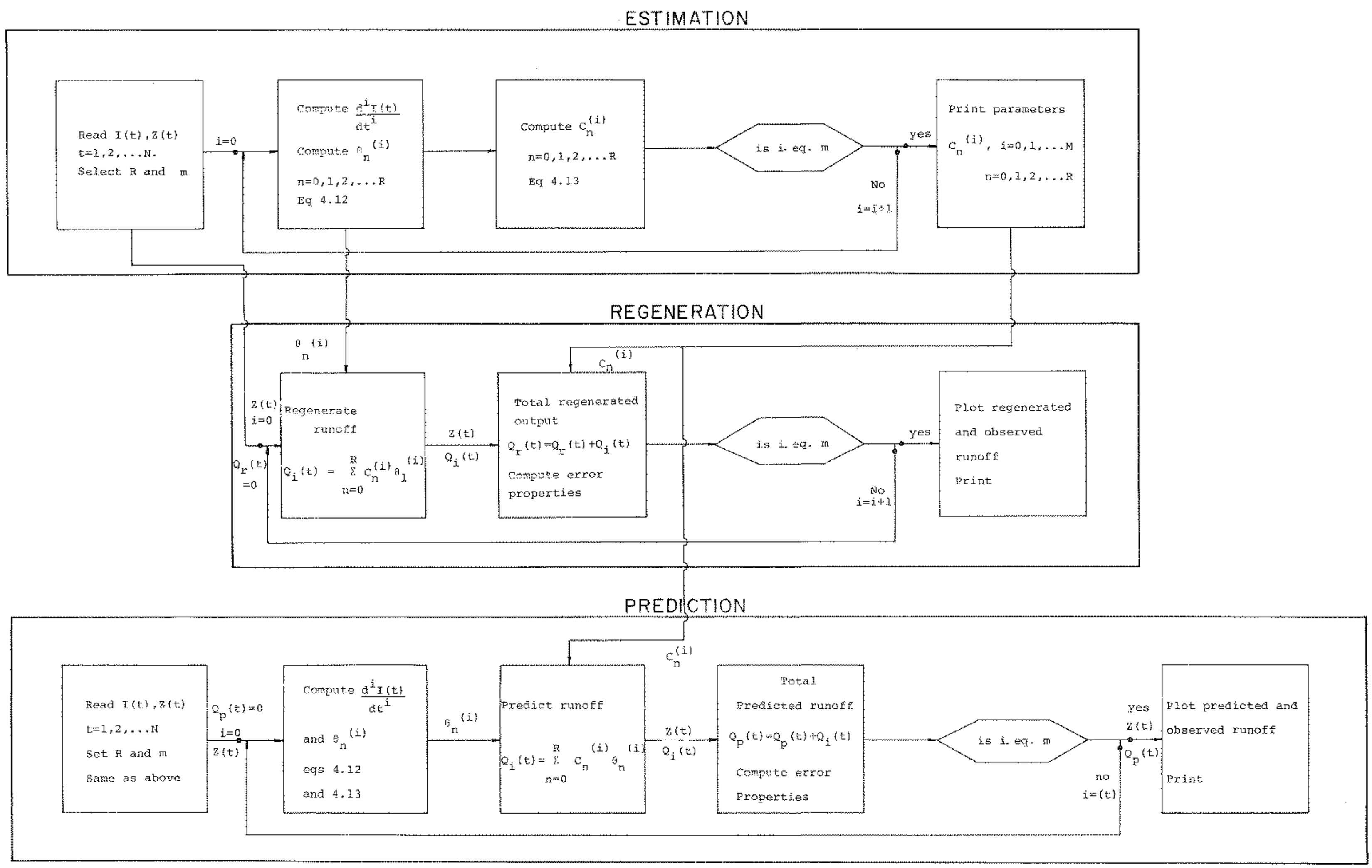
39. Zadeh, L. A., "Optimum Nonlinear Filters," J. Appl. Physics, Vol. 24, No. 24, 1953.

APPENDIX

The flow charts which clarify the computational sequences for the two models discussed in the text are given in the appendix. The computational sequence for the functional series model is given in flow chart A1, and in the flow chart A2, the computational sequence for the nonlinear stochastic model is outlined.



A1: COMPUTATIONAL SEQUENCE FOR THE FUNCTIONAL SERIES MODEL



A2: COMPUTATIONAL SEQUENCE FOR THE NONLINEAR STOCHASTIC MODEL